
인터넷 트래픽 특성을 이용한 ATM 망의 최적설계에 관한 연구

최삼길* · 김동일*

A Study on the optimization design of ATM network Using Internet Traffic Characteristics

Sam-gil Choi* · Dong-il Kim*

요 약

기존의 큐잉분석은 네트워크 용량설계 및 성능 예측시 상당히 유용하지만, 실제의 경우 큐잉 분석으로부터 예측된 결과는 실제 관측된 성능과는 상당히 다르다. 또한 LAN, WAN 및 VBR 트래픽 특성에 관한 최근의 실험적 연구들은 기존의 Poisson 가정에 의한 모델들이 네트워크 트래픽의 장기간 의존성 및 self-similar 특성을 과소평가 함으로써, 실제 트래픽의 특성을 제대로 나타낼 수 없다는 것을 지적해 왔다. 본 논문에서는 exactly self-similar 모델링인 Pareto-like ON/OFF 소스 모델링에 의한 트래픽과 기존의 Poisson 모델링에 의한 트래픽을 비교 분석함으로써, self-similar 트래픽이 실제의 Ethernet 트래픽 특성을 잘 반영한다는 것을 보임으로써 최적화된 ATM 네트워크 설계 및 성능분석을 이루고자 한다.

ABSTRACT

Traditional queueing analyses are very useful for designing a network's capacity and predicting their performances, however most of the predicted results from the queueing analyses are quite different from the realistic measured performance. And recent empirical studies on LAN, WAN, and VBR traffic characteristic have indicated that the models used in the traditional Poisson assumption cannot properly predict the real traffic properties due to underestimation of the long-range dependence of network traffics and self-similar properties. In this paper, It is also shown that the self-similar traffic reflects real Ethernet traffic characteristics by comparing Pareto-like ON/OFF source model which is exactly self-similar model to the traditional Poisson model. It is also performed optimization design and performance analysis of ATM network using Internet traffic characteristics.

키워드

Pareto-like on/off 소스모델링, 최적설계, 자기유사특성

1. 서 론

멀티미디어 서비스는 80년대 초 PC, LAN 그리고 디지털 PBX로 시작하여 최근의 기하급수적으로 증가하는 인터넷 사용자와 고성능 컴퓨터의 대중화 그리고 고속 사설 망의 증가등 급격한 시장 변화는 단

일 서비스 위주의 통신 제품들의 수요 한계와 함께 새로운 개념의 통신제품의 필요성을 예측 가능하게 하고 있다. 이러한 추세 속에 80년대 후반부터 ATM 기술을 기반으로 활발히 연구되어온 새로운 네트워크 기술은 멀티미디어 서비스라는 명제를 실현하기 위한 네트워크 구성체계를 요구하고 새로운 트래픽

* 동의대학교 정보통신과

모델의 분석기법을 도입하여 네트워크의 지연특성과 처리율 분석을 가져오게 되었다.

기존의 데이터 트래픽에 관련된 일반적인 가설은 많은 수의 독립적인 트래픽 스트림들을 멀티플렉싱하는 것이 포아송(Poisson) 과정으로 귀착된다는 것이다. 기존의 이러한 가정과 결과적인 큐잉분석이 초창기의 ATM 스위치를 매우 적은 버퍼(10~100cell)를 가진 스위치를 생산하도록 하였고, 실제의 트래픽을 수용했을 때 엄청난 셀 손실들이 발생되었고 그 스위치들을 다시 설계하는 결과를 초래하게 되었다. 그러므로, 네트워크의 설계 및 성능평가에 앞서 그 네트워크의 트래픽 특성을 이해하는 것이 매우 중요하다 하겠다. 네트워크 트래픽에 관한 최근의 논문들은 LAN, WAN 및 VBR 비디오 트래픽이 통계학적으로 자기유사 과정에 의해 더욱 더 잘 모델링 된다는 것을 설득력 있게 주장해 왔다[1][2]. 따라서 본 논문에서는 이러한 비결정적이고, 불규칙적인 트래픽의 버스트 특성에서 규칙성의 패턴을 추출하기 위해 시계열의 확대나 이동에 상관없이 통계적 특성(평균, 분산, 자기상관함수)을 유지하는 자기유사성을 정의하고, 자기유사 특성을 갖는 파레토(Pareto) 분포 모델과 기존의 네트워크 성능분석법인 지수 분포 모델을 ATM 시뮬레이션 네트워크에 인가하여 망에서 일어나는 현상을 비교·분석함으로써 최적화된 네트워크 설계 요소 및 성능변수를 구하고자 한다.

II. 자기유사 특성

1. 자기유사성의 정의

일반적으로, 결정적이고 주기적 신호는 시간이동에 관하여 불변인 특징이 있다. 그 신호는 시간상으로 여러 주기가 이동하더라도 동일한 신호이다. 예를 들어, 주기가 T인 결정적이고 주기적인 함수 g(t)는 식 (1)과 같다[3][4][7].

$$g(t) = g(t+aT), a = 0, \pm 1, \pm 2, \dots \quad (1)$$

이와 대조적으로, 정상확률과정(stationary stochastic process)에 대해서는 그 과정의 통계는 시간 이동에 불변이다. 평균은 시간에 독립적이고 자기상관함수는 단지 시간의 차이에만 의존한다. 결정적인 자

기유사 신호는 스케일의 변화에 무관하다. 즉, 확률 과정에 대해서 그 과정의 통계가 시간 스케일의 변화에 따라 변화하지 않는다. 정상적이고 정량적인 측면에서, 그 과정은 스케일의 특징이 없다. 비결정적인 자기유사성은 자연적, 인공적 현상에서 매우 일반적이다.

2. 연속시간 정의

어떠한 실수 $a > 0$ 에 대해, 확률과정 $a^{-H}x(at)$ 가 $x(t)$ 와 통계적으로 동일한 특성을 가진다면, 확률과정 $x(t)$ 는 파라미터 $H(0.5 \leq H \leq 1)$ 를 가지고 통계적으로 자기유사하다. 이러한 관계는 다음의 3가지 조건으로 표현된다[5][6][7].

$$E[x(t)] = E[x(at)] \quad \text{Mean}$$

$$Var[x(t)] = \frac{Var[x(at)]}{a^{2H}} \quad \text{Variance}$$

$$R_x(t, s) = \frac{R_x(at, as)}{a^{2H}} \quad \text{Autocorrelation}$$

Hurst 또는 self-similarity 파라미터 H는 self-similarity의 핵심척도이다. 다시 말하면, H는 통계적인 현상의 지속성에 대한 척도이고 확률과정의 장기간 종속에 대한 척도이다. H=0.5의 값은 self-similarity의 부재를 나타내고, H가 1에 가까울수록, 지속성의 정도 또는 장기간의 종속의 정도는 더욱 커진다.

3. 이산시간 정의

정상 시계열(stationary time series) x에 대해, m-집합된 시계열 $x^{(m)} = \{x_k^{(m)}, k=0, 1, 2, \dots\}$ 를 인접한 m 크기의 블록을 겹침 없이(nonoverlapping) 원래의 시계열을 합계함으로써 정의한다. 이것은 식 (2)와 같이 표현된다.

$$x^{(m)} = \frac{1}{m} \sum_{i=km-(m-1)}^{km} x_i \quad (2)$$

예를 들어, x_k 는 식 (3)과 같이 정의된다.

$$x_k^{(3)} = \frac{x_{3k-2} + x_{3k-1} + x_{3k}}{3} \quad (3)$$

집합된 시계열을 관찰하는 한 방법은 타임스케일을 압축하는 기법과 같다. 즉, $x^{(1)}$ 는 이러한 시계열의 최대 해상도이고, $x^{(3)}$ 는 비율 3으로 축소된 것이다. 즉, 세 시점의 샘플에 대해 평균을 구함으로써, 최상의 확대도에서 얻을 수 있는 우수한 상세도를 잃게 된다. 만약 이것의 확률과정의 통계(평균, 분산, 자기상관함수)가 동일한 압축 사본을 간직하고 있다면 자기유사 과정으로 다룰 수 있다. $x^{(m)}$ 의 에르고딕 과정에 대해, 시간평균은 조화평균과 동일하고, 시간평균의 분산은 조화평균과 동일하다. 만약 시간평균의 분산이 m 이 매우 커짐에 따라 영으로 수렴하게 되면, 이것은 자기유사 과정이 아니다. 만약 확률과정 x 가 모든 $m=1, 2, \dots$ 에 대해서 식(4), (5)와 같다면 파라미터 β ($0 < \beta < 1$)에 대하여 정확하게 자기유사하다고 한다.

$$Var(x^{(m)}) = \frac{Var(x)}{m^\beta} \quad (4)$$

$$R_{x^{(m)}}(k) = R_x(k) \quad (5)$$

파라미터 β 는 앞에 정의했던 허스트 파라미터 ($H=1-(\beta/2)$)와 관련이 있다. 정상과정 및 에르고딕 과정에서는 $\beta=1$ 이고 시간 평균의 분산은 $1/m$ 비율로 감소하게 되지만, 자기유사 과정에서는 시간평균의 분산은 더욱 천천히 ($1/m^\beta$) 감소하게 된다. 확률과정 x 가 충분히 큰 모든 k 에 대해 식(6), (7)과 같다면 점근적으로(asymptotically) 자기유사하다고 한다.

$$Var(x^{(m)}) = \frac{Var(x)}{m^\beta} \quad (6)$$

$$R_{x^{(m)}}(k) \rightarrow R_x(k), \text{ as } m \rightarrow \infty \quad (7)$$

따라서, 이러한 자기유사성의 정의에 의해서, 집합된 과정의 자기상관은 원래의 확률과정과 동일한 형태를 가진다. 이것은 변이성 또는 버스트한 정도가 서로 다른 타임 스케일에서 동일하게 나타난다는 것을 의미한다. 앞의 정의에서 흥미로운 것은 집합된

자기유사 과정의 자기상관이 $m \rightarrow \infty$ 함에 따라 영으로 수렴하지 않는다는 것이다. 이것은 기존의 패킷 데이터 모델에 사용되는 확률과정과는 대조를 이루고, 식 (8)과 같은 관계를 가진다.

$$R^{(m)}(\tau) \rightarrow 0 \text{ as } m \rightarrow \infty \quad (8)$$

$R(\tau)=0$ 인 자기상관 함수는 백색잡음과 일치한다. 즉, 균일 전력 스펙트럼을 가지는 순수한 랜덤확률과정이다. 그 다음 흥미로운 특징은 $x(m)$ 의 분산이 $m \rightarrow \infty$ 가 됨에 따라 $1/m$ 보다 더욱 천천히 감소한다는 것이다. 즉, $1/m^\beta$ 에 비례해서 감소한다. 이에 반해, 기존의 패킷 트래픽 모델에 사용되는 확률과정에서는 분산이 $1/m$ 으로 감소하는데, 에르고딕 과정에 대한 이런 감소율은 $x(m)$ 가 시간평균이고 표본크기 (sample size) m 을 갖는 표본평균과 통계학적으로 동등하다는 것을 쉽게 알 수 있다[10][11].

III. ATM 시뮬레이션 네트워크 트래픽 분석

본 논문에서는 그림 1과 같은 ATM 시뮬레이션 네트워크를 구성하고, 각 노드에서 heavy-tailed 특성을 가지는 대표적인 트래픽 분포 모델인 파레토 도착 분포를 갖는 트래픽을 발생시켰을 때, shape 변수 α ($1 \leq \alpha \leq 2$)의 값에 따른 시간별 트래픽의 상태와 drop율, 네트워크 이용률, 그리고 스위치의 입력버퍼 변화율을 분석한다. 우선, 각 노드에 입력되는 트래픽 소스의 큐 도착 분포변수를 파레토 소스 모델로 설정하고, α 의 값을 변화시켰을 때, 스위치 노드에 입력되는 중첩된 트래픽들의 자기유사 정도가 입력변수인 α 의 값에 따라 이론적인 허스트 값을 만족하는지 살펴보았다.

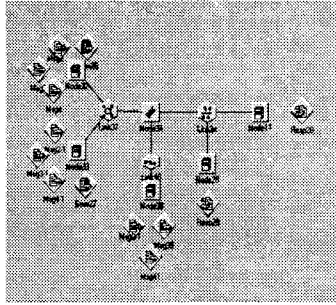


그림 1. 시뮬레이션 네트워크 모델
Fig. 1. Simulation Network model

$$\alpha = 3 - 2H \quad (9)$$

식(9)에 의해, $\alpha = 2$ 가 일 때 $H = 0.5$ 가 되고, $\alpha = 1.1$ 일 때 $H = 0.95$ 가 되어 비 자기유사성의 트래픽 모델과 자기유사성의 트래픽 모델을 만들게 되고, 스위치의 입력버퍼에서 도착하는 트래픽의 버퍼사용율 트레이스를 관찰하고 자기유사성을 분석하였다. 자기유사성을 분석하는 방법에는 여러 가지 방법이 있지만 본 논문에선 가장 간단한 Variance-time plot 방법을 사용하였다. 그림 2, 3은 α 값에 따른 트래픽 트레이스를 나타낸 것이고 그림 4, 5는 α 값에 따른 Variance-time plot을 사용하여 H값을 분석한 것이다[8][9].

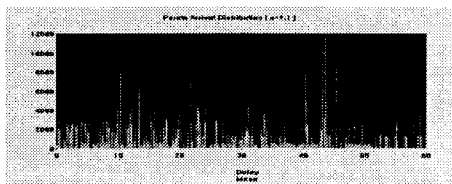


그림 2. 파레토 도착분포 트레이스 ($\alpha=1.1$)
Fig. 2. Pareto arrive distribution trace ($\alpha=1.1$)

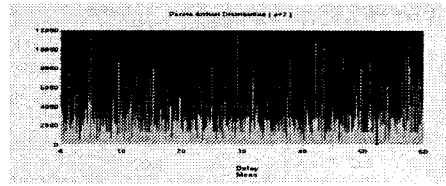


그림 3. 파레토 도착분포 트레이스 ($\alpha=2$)
Fig. 3. Pareto arrive distribution trace ($\alpha=2$)

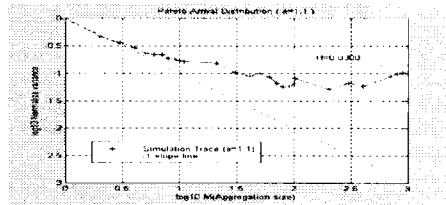


그림 4. 파레토분포 ($\alpha=1.1$)의 Variance-time plot
Fig. 4. Variance-time plot of Pareto distribution ($\alpha=1.1$)

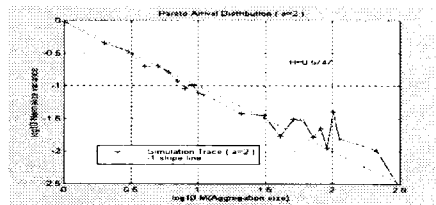


그림 5. 파레토분포 ($\alpha=2$)의 Variance-time plot
Fig. 5. Variance-time plot of Pareto distribution ($\alpha=2$)

소스 트래픽의 파레토 변수에 따른 자기유사 정도를 관찰한 결과, 파레토 shape 변수 α 가 1.1 일 경우, 이론적인 H 의 값은 0.95가 된다. ATM 시뮬레이션 네트워크에서 다음과 같은 변수로 지정 후 트래픽을 인가한 결과 H 는 0.8380의 값이 되어 이론적인 값보다는 작지만, 자기유사 트래픽 특성을 나타내고 있다. 그리고 $\alpha=2$ 인 경우 $H=0.5747$ 로서 포아송 트래픽과 거의 유사한 H값을 나타내고 있다.

다음은 이러한 트래픽 소스를 ATM 시뮬레이션 네트워크 모델에 인가하였을 때, 네트워크에 일어나

는 통계적인 현상을 분석한다.

표 1. Drop을 비교

Shape variable number(α)	1.1	1.3	1.5	2
Drop Packet(number)	9	7	21	24

첫 번째로, drop율을 살펴보면 파레토 ($\alpha = 1.1$)는 입력 큐에 도착하는 트래픽의 heavy-tail 특성에 의해 버퍼 도착확률이 분산되는 반면, SRD 특성을 갖는 파레토 ($\alpha = 2$)는 단기간 버퍼에 집중적으로 도착하기 때문에 중첩된 트래픽에서는 drop의 확률이 높아지게 된다.

표 2. 지연을 비교

Shape variable number(α)	1.1	1.3	1.5	2
Average delay(ms)	4.774	5.0046	4.951	5.2744

두 번째로, ATM 시뮬레이션 네트워크에서 지연율을 비교해 보면, 버퍼의 오버플로의 확률이 높은 파레토 ($\alpha = 2$)에 의해 버퍼에서 서비스를 위해 대기하는 시간이 길어지고, 따라서 전체 시스템의 지연에 영향을 미치게 된다.

마지막으로, 채널 이용률은 shape 변수의 증가에 따라, 각 링크의 이용률이 증가함을 보이고 있다.

이러한 자기유사의 heay-tailed 특성을 나타내는 파레토 모델을 네트워크에 적용함으로써, 기존의 포아송 도착패턴에 의한 지수 분포모델의 트래픽 특성이 실제 트래픽의 통계적인 특성과는 명백한 차이가 있고, 따라서 자기유사 특성을 고려한 네트워크 분석이 요구된다.

표 3. 채널 이용률 비교

Shape variable number(α)		1.1	1.3	1.5	2
Channel Utilization (%)	Ethernet	2.742	3.721	5.112	7.164
	Token Ring	0.824	0.915	1.562	1.972

IV. 네트워크 최적 설계 요소 검출

지금까지는 네트워크 트래픽의 이론적인 분석을 하였다. 사후 조사 방법으로 실제 네트워크 트래픽을 장기간 관찰하였고, 이러한 결과를 얻기 위해서는 시스템의 상태정보를 수집하여 상황에 따른 시스템 환경을 변경시킬 수 있는 트래픽 산출 및 분석 시스템 개발이 필요하다. 이러한 네트워크 성능 관리 시스템을 통하여 전체 네트워크의 효율성을 이루고 체계적인 성능 관리를 부여하여 관리자에게 분석 도구로서의 편의성을 제공함은 물론, 효율적인 네트워크 설계를 통한 안정적인 네트워크 서비스 제공이 가능하게 된다.

본 논문에서는 장기간 트래픽 산출 및 분석을 위하여 Linux Platform에 SNMP 와 MIB-II를 이용하는 MRTG(Multi Router Traffic Grapher)를 구동하여 동의대학내 망의 게이트웨이와 외부 ISP와 연결된 라우터 인터페이스를 모니터링 함으로써 인터페이스별 트래픽 이용량을 알 수 있으며, 네트워크 선로의 혼잡 발생 가능성을 추측할 수 있고, 또한 방송형 트래픽과 비방송형 트래픽을 비교 분석함으로써 네트워크 성능에 좋지 않은 영향을 미치는 네트워크 관리요소를 분별해 낼 수 있다. 따라서 네트워크상의 장기적인 성능 관리와 대역폭 계획을 함으로써 최적의 네트워크 상태를 유지하게 된다. 그림 6은 게이트웨이 라우터(134.75.226.1)의 인터페이스별 이용률 현황을 보이고 있다. 좌측 상단의 그래프는 학내 망 내부와 접속되어 있는 이더넷 인터페이스(10Mbps)로

송신이 수신 요구보다 높다. 이는 게이트웨이를 벗어나 인터넷 접속 요구가 높음을 알 수 있다.

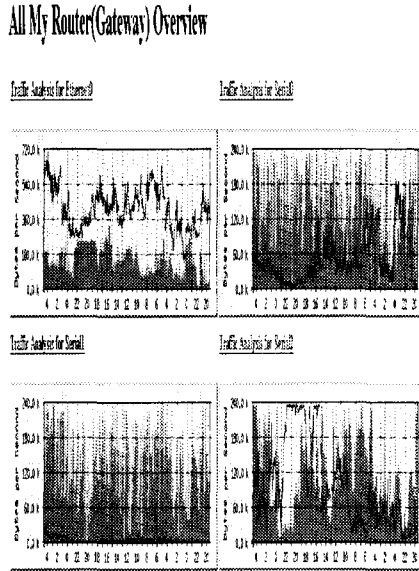


그림 6. 게이트웨이 라우터의 각 인터페이스 이용률
Fig. 6 Each interface utilization rate of gateway router

나머지 직렬 라인으로 표시되어 있는 그래프는 외부 ISP(Internet Service Provider)와 학교 네트워크와 연결되어 있으며, 데이터의 송신(uplink) 보다는 수신(downlink)의 비가 높음을 보이고 있다.

S0(203.241.174.22)와 S1(203.241.174.26)은 연구전산망과 E1(2.048Mbps) 회선으로 각각 연결되어 있으며, S2(172.19.13.253)는 시스템 공학연구소에 연구전산망과 같은 E1회선으로 연결되어 있다.

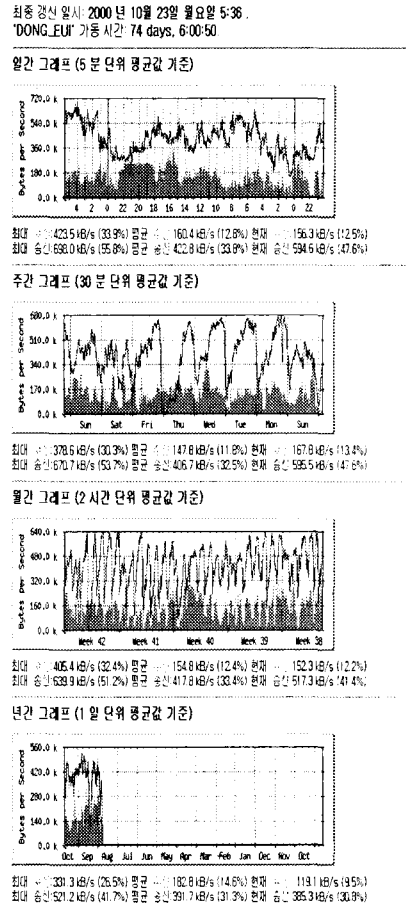


그림 7. 게이트웨이 인터페이스의 시간단위 이용률
Fig. 7. Time unit utilization rate of gateway interface

그림 7은 게이트웨이 라우터의 인터넷 인터페이스로 내부망 이용자들의 시간단위별 이용률을 보이고 있다. 수신보다는 송신의 비가 높은 결과를 보이고 있고, 이러한 결과는 외부망(인터넷)으로 나가기 위한 내부망 사용자들을 송신 이용률이 높음을 나타낸다.

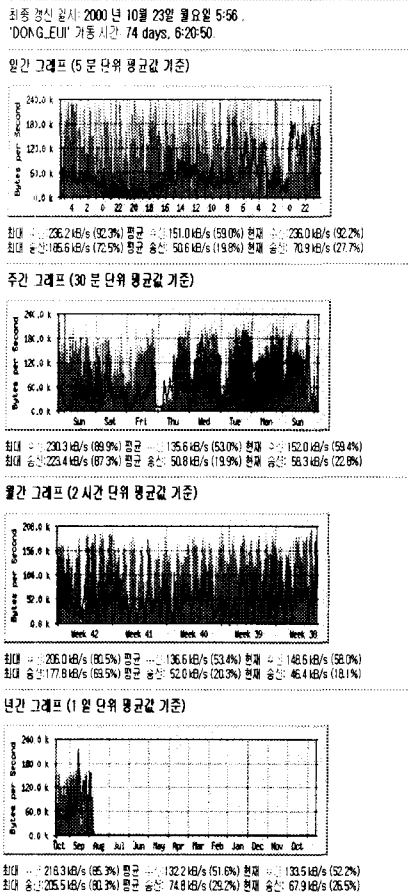


그림 8. 게이트웨이 인터페이스(203.241.174.22)의 시간단위 이용률
 Fig. 8. Time unit utilization rate of gateway interface (203.241.174.22)

그림 8은 게이트웨이 라우터의 직렬 인터페이스(S0)의 이용률을 보이고 있고 송신 이용률(uplink) 보다는 수신 이용률(downlink)이 높아 인터넷으로부터 다양한 웹 콘텐츠를 로딩하거나, 메일확인, 파일 송·수신 작업이 활발히 일어나고 있음을 보이고 있다. 상위 그래프들은 총 74일간의 인터페이스 이용률을 모니터링 했으며, 일일 이용률 또는 주 단위의 이용률 형태가 비지(busy)와 유희(idle) 시간의 일정한 패턴으로 나타난다. 최상위의 그래프는 5분 단위로 게이트웨이 라우터의 정보를 표시함으로써 원격지의

네트워크 관리자는 실시간 네트워크 상태를 감시하거나, 제어를 통해 문제점을 분석하고 해결할 수 있는 능력을 갖게 된다. 또한 이용률 트래픽의 장기간 모니터링의 결과는 실제 네트워크에 필요한 회선 용량, 버퍼사이즈, 정체의 원인 그리고 적절하지 못한 부하의 분산 정책 등을 수정 가능하게 하며, 네트워크상의 장기적인 성능 관리와 대역폭 계획을 함으로써 최적의 네트워크 상태를 유지하게 된다.

V. 결론

이제까지 네트워크 트래픽의 특성 분석을 위해 이론적인 접근과 최근 관심과 많은 연구가 되어지고 있는 자기유사 특성에 대해 언급하였다. 새로운 시스템을 구성하거나, 또는 미래 예측을 통한 시스템 용량과 안정된 서비스를 보장하기 위해 지금까지 포아손 입력 패턴을 가정한 큐잉 모델을 보편적으로 사용하여 왔다. 그러나 최근의 연구 결과는 그러한 가정이 트래픽의 특성을 반영하지 못함을 지적하였고, 새로운 트래픽 특성인 자기유사의 개념을 도입한 큐잉 분석이 필요함을 제기하고 있다. 따라서 본 논문에서는 트래픽의 자기유사 개념을 정의하고, 허스트 변수의 값이 1에 가깝도록 인위적인 자기유사 트래픽을 만들어 ATM 시뮬레이션 네트워크에 입력했을 때 네트워크의 반응을 분석하였다. 이러한 결과로서 기존의 포아손 트래픽 모델로는 단기간 의존성 기반의 트래픽 모델(포아손, 지수분포)으로써 평균 패킷 전송 지연이나 큐의 크기를 설계한 네트워크는 어떠한 시점에서 트래픽의 과도한 증가로 인한 큐의 정체 또는 폭주가 발생하게 되었을 때 장기간 의존성 특성을 과소 평가 함으로 인해 큐에서는 이러한 트래픽을 처리하지 못하고, 입력 패킷의 폐기 또는 재전송 요구로 인해 결과적으로 망 전체의 트래픽을 다시 증가시키는 결과를 초래하게 된다는 것을 알 수 있었다. 또한, 장기간 네트워크 상태를 모니터링 하여 얻은 결과로써 성능 개선을 위한 최적의 설계 요소를 검출하였다.

참 고 문 헌

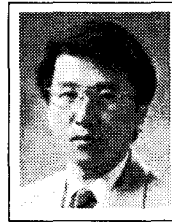
저 자 소 개

- [1] W.Willinger, M.S.Taqqu, R. Sherman, "Self-Similarity Through High Var-iability", ACM/Sigcomm, 1995.
- [2] Schroeder, M., "Fractal, Chaos, Power Laws : Minutes from an Infinite Paradise", Freeman, 1991.
- [3] William Stallings, "High-Speed Networks", Prent-iceHall, pp. 181-208, 1999.
- [4] J. M. Pitts, J. A. Schormans, "Introduction to ATM design and performance: with applications analysis software", John Wiley & Sons Ltd., 1996.
- [5] Peyton Z. Peebles, Jr. Probability, "Random Variables, and Random Signal Principles", McGraw-Hill, pp. 134-198, 1993.
- [6] K.Meirer-Hellstern, P.Wirth, Y-L. Yan and D.Hoeflin, "Traffic Models for ISDN Data Users : Office Aut-omation Application", Proc. ITC-13, Copenhagen, 1991.
- [7] A. Erramilli and W. Willinger, "Fractal Prop-erties in Packet Traffic Measurement", Proc. of ITC Regional Seminar, St. Petersburg, 1993.
- [8] J.Beran, R. Sherman, M.S. Taqqu and W. Willi-nger, "Long-Range Dependence in Variable Bit Rate Video Traffic", IEEE Tracsaction on Communications 43, No.2/3/4, pp. 1566-1579, 1995.
- [9] D.Wilson, "Self-similar Traffic Modeling for Highspeed Networks", ConneXions, Nov, 1994.
- [10] P., Pruthi, "Heavy-tailed ON/OFF Source Beha-vior and Self-Similar Traffic", ICC' 95, 1995.
- [11] Norman S.Matloff, "Probability Modeling and Computer Simulation: Applied to Engineering and Computer Science", PWS-Kent, pp.97-107, 1988.



최삼길(Choi Sam Gil)

2000년 11월 제4권 제4호참조
2002년 6월 현재 동의대학교
정보통신공학과 교수



김동일(Kim Dong Il)

2000년 6월 제4권 제2호 참조
2002년 6월 현재 동의대학교
정보통신공학과 교수