

BGP 밸리-프리 라우팅 정책에 기반한 국내 AS 레벨 인터넷 토폴로지의 파워-로 지수[☆]

Power-Laws Exponents of the Domestic AS-level Internet Topology based on the Valley-free BGP Routing Policy

강 구 흥*
KooHong Kang

요 약

Faloutsos et al.(1,2)은 파워-로(power-law)를 이용해 노드 차수와 같은 Autonomous System (AS) 레벨 인터넷 토폴로지 특성의 중꼬리(heavy-tailed) 분포를 성공적으로 나타내었다. 이러한 결과는 파워-로 지수(exponents)를 이용해 이들 인터넷 토폴로지 특성을 간단 명료하게 나타낼 수 있게 한다. 본 논문에서는 BGP 밸리-프리 라우팅 정책에 기반한 국내 AS 레벨 인터넷 토폴로지 속성 - 노드 차수, 홉 수에 따른 노드 쌍의 수, 그리고 그래프의 고유치 - 의 파워-로 지수를 조사하였다. UCLA IRL 연구실이 제공하는 실제 AS 레벨 데이터 셋을 이용하였으며 이들 파워-로 근사는 상관계수(correlation coefficient)가 각각 90.7%, 96.5%, 그리고 97%로 조사되었다. 특히, 실질적인 AS 레벨 토폴로지 직경이 3홉 이내에 존재하는 AS 노드 쌍이 전체 91% 이상이며, 따라서 국내 AS 레벨 토폴로지가 비교적 잘 정리된 것으로 판단된다.

ABSTRACT

Faloutsos et al.(1,2) showed successfully that the power-laws describe the heavy-tailed distributions of the AS level Internet topology properties such as node degree. This result allows us to represent the characteristics of AS-level Internet topology using some power-law exponents with elegant and simple. In this paper, we obtained the power-law exponents of the domestic AS-level Internet topology properties - the node degree, the number of pairs within hops, and eigenvalues of the graph - based on the valley-free BGP routing policy. We used the real data sets from UCLA IRL laboratory, and showed that these power-laws fit the real data pretty well resulting in correlation coefficient of 90.7%, 96.5%, and 97%, respectively. In particular, rounding the effective diameter to three, approximately 91% of the pairs of nodes are within this distance; that is, we might conclude our topology is pretty well organized.

☞ KeyWords : AS-level Internet Topology, Power-law Exponent, BGP, AS 레벨 인터넷 토폴로지, 파워-로 지수, BGP

1. 서 론

1.1 연구 배경

컴퓨터의 대중화와 정보화 사회의 도래와 함께 인터넷의 속도와 복잡도는 놀라운 속도로 증가하고 있다. 인터넷은 다양한 네트워크 및 인터넷 서비스 제공자(NSP, ISP: Network (or Internet) Service Provider)에 의해 관리되고 운영되는 수 천 개의 AS(Autonomous System)로 연결되어 있으며 BGP (Border Gateway Protocol) 라우팅 프로토콜에 의해 이들 AS 간 라우팅 패스가 결정된다[3,4]. 따

* 정 회 원 : 서원대학교 정보통신공학과 부교수
khkang@seowon.ac.kr

[2010/05/06 투고 - 2010/05/10 심사 - 2010/07/21 심사완료]

☆ The author wishes to thank the member of Internet Research Lab, Computer Science Department in UCLA, for the Internet Topology Collection. In particular, the author would like to thank Dr. Ricardo V. Oliveira in UCLA for his fruitful discussions about the data sets. The author also would like to thank Prof. Heeman Lee in Seowon University for the initial version of BGP routing path building program used in this paper.

라서 인터넷은 각 AS가 노드, 그리고 두 AS 사이 BGP 피어링(peering)이 링크로 표현된 하나의 AS 레벨 토폴로지 그래프로 나타낼 수 있다. 지난 수 년 간 많은 사람들은 인터넷이 어떤 형태로 상호 연결되어 있으며 어떤 토폴로지 특성을 갖고 있는지에 관한 연구를 진행해 오고 있다[1-7]. 이러한 AS 레벨 토폴로지 연구는 인터넷 운영과 인터넷 기술 연구에 있어 매우 중요한 의미를 가진다. 즉 AS 레벨 토폴로지 특징을 이용해 보다 효율적인 프로토콜을 설계할 수 있으며, 모의실험을 위해 보다 정확한 모델을 개발할 수 있다. 뿐만 아니라 프로토콜 분석과 미래 인터넷 토폴로지를 예측하기 위한 다양한 토폴로지 관련 파라미터를 예측할 수 있게 한다.

AS 레벨 토폴로지를 생성하는 방법은 크게 두 가지가 있다. 첫 번째 방법은 일반인들에게 공개된 RouteViews와 RIPE-RIS가 제공하는 데이터 셋을 이용해 직접 AS 레벨 토폴로지를 생성하는 방법(이하, public-view라고 약칭)과 두 번째 방법으로 AS 레벨 토폴로지를 자동으로 생성하는 발생기를 이용할 수 있다. RouteViews와 RIPE-RIS와 같은 프로젝트들은 현재 운영 중인 수 백 개의 라우터와 BGP 세션을 형성해 데이터를 수집하고 있으며, 일반인들이 접근해 이들 데이터 소스를 사용할 수 있도록 공개되고 있다[9-11]. 그러나 최근 참고문헌[5]와 [6]에서는 이들 공개 데이터 소스만을 활용해 인터넷 AS 토폴로지를 추론할 경우, 전체 AS 토폴로지의 상당 부분이 발견되지 못하는 문제점을 지적하고, 새로운 발견적 방법(heuristic method)를 이용하여 관찰된 AS 토폴로지로부터 모든 사용자-제공자(customer-provider) 링크를 정확하게 확인할 수 있는 방법을 제안하였다. 한편, AS 레벨 인터넷 토폴로지 발생기 Inet-3.0[12]과 NIT[13]는 현 인터넷 토폴로지 특징을 흉내 내어 AS 노드 수와 노드 차수(degree) 정보를 입력 받아 AS 토폴로지 연결 정보를 자동으로 생성해 낸다. 그러나 이들 발생기들이 현 인터넷 토폴로지 특징을 정확하게 반영하기에는 아직

여러 가지 문제점들이 있다.

앞에서 언급한 바와 같이 인터넷 AS 레벨 토폴로지에 관한 연구가 상당히 많이 이루어지고 있으나, 국내 연구 결과물은 거의 찾아보기 불가능하며 한국인터넷정보센터(KRNIC: Korea Network Information Center)에서 관리하는 국내 AS 번호 정보를 확인할 수 있는 수준이다. 이러한 현상은 국내 인터넷 연구가 대부분 라우터 레벨에서 이루어지고 있으며 AS 레벨 토폴로지 연구의 중요성을 아직 인식하지 못하고 있기 때문이다. 뿐만 아니라 국내 인터넷 시장은 몇몇 거대 기업이 네트워크 사업 및 인터넷 사업을 병행하고 있어 NSP와 ISP의 시장 경계가 모호하며, 따라서 실질적인 AS 레벨 토폴로지의 노출에 따른 이들 기업 정책이 일반인들에게 공개되는 것을 반기지 않을 수 있다.

1.2 연구 동기

Faloutsos et al. [1]는 AS 레벨 인터넷 토폴로지 그래프 상에서 노드 차수 분포 및 h 홉 내 전체 AS 노드 쌍 분포, 그리고 그래프의 고유치(eigenvalue) 분포가 ‘파워-로(power-law)’ 특징을 갖는다는 연구 결과를 최초로 발표하였다. 이러한 연구결과는 당시 네트워크 관련 분야에서 전혀 예상하지 못한 사실이었으며, 이러한 현상을 만들어 내는 원인 파악에 많은 노력을 기울여 왔다 [3,8]. 본 논문에서는 국내 인터넷 토폴로지 특징을 AS를 기준으로 이러한 파워-로 특징을 분석한다. 이러한 분석 결과는 앞에서 설명한 바와 같이 국내 인터넷 운영 및 연구에 중요한 기초 자료가 될 것이다.

외국의 많은 AS 레벨 인터넷 토폴로지 연구가 물리적 연결 구조 특징을 밝히는데 머물러 있다. 그러나 BGP 라우팅 프로토콜에 의해 실질적인 AS 레벨 경로가 결정되며, 따라서 본 논문에서는 기본적인 BGP 라우팅 정책을 기반으로 한 인터넷 AS 레벨 라우팅 패스 토폴로지(이하, 본 논문에서는 간단히 ‘AS 레벨 토폴로지’로 약칭)의 다

양한 특징을 파악한다. BGP 라우팅 프로토콜의 가장 큰 특징은 각 AS가 자신의 관리정책에 의해 최적 경로를 선택하고 경로정보를 송·수신하는 것이다[4]. 그러나 수 천 개의 각 AS 별 고유한 관리정책을 모두 반영하여 AS 레벨 토폴로지 특징을 파악하는 것은 사실상 불가능한 것이며 또한 큰 의미를 가지지 못한다. 따라서 본 연구에서는 BGP ‘밸리 프리(valley-free)’기반 정책을 기반으로 미국 UCLA 전자계산학과 Internet Research Lab (IRL)[11] 에서 제공하는 실제 인터넷에서 추출된 AS 레벨 토폴로지 자료를 활용하여 국내 AS 들의 파워-로 특징들을 알아 본다. 이제부터 인터넷 AS 레벨 토폴로지는 국내 721개(2009년12월2일 기준, KRNIC(Korea Network Information Center) 자료 [14]) AS를 대상으로 한정한다.

서론에 이어, 제2장에서는 AS 레벨 인터넷 토폴로지의 파워-로 관련 기존 연구 결과들을 간략히 기술하고 본 논문에서 사용할 BGP ‘밸리-프리’ 라우팅 정책에 대해 설명한다. 제3장에서는 국내 AS 레벨 인터넷 토폴로지의 생성 방법과 제2장에서 설명한 파워-로 특성을 기준으로 국내 인터넷 토폴로지 특성을 알아본다. 마지막으로, 제4장에서 결론 및 향후 연구 방향에 대해 기술하였다.

2. AS 레벨 인터넷 토폴로지

2.1 관련 연구

Faloutsos et al.[1]은 1997년과 1998년 사이 AS 레벨 토폴로지의 vertex 차수 분포가 파워-로 특성을 보임을 최초로 밝혔다. 그러나 Chen et al. [8]은 참고문헌[1]에서 사용된 데이터 소스가 20-50% 혹은 그 이상의 AS 노드 간 링크가 빠져 있음을 지적하고 노드 차수 분포가 중꼬리 (heavy-tailed) 분포이기는 하나 정확하게 파워-로 특성을 따르지는 않는다고 지적하였다. 이후, Siganos et al.[2]은 2003년 다시 5년간 수집된 보다 완전한 인터넷 토폴로지를 대상으로 상관계수 (correlation coefficient)가 99% 이상인 파워-로 특

성을 재확인하여 발표하였다. 한편, Oliveira[5] 역시 public view 만을 사용할 경우 전체 AS의 4%만이 전체 링크 연결구성을 확인할 수 있음을 그들의 연구결과를 통해 보여주고 있다. 나머지 96% AS들은 일부 사용자-제공자 링크만을 나타내고 대부분의 피어(peer) 링크는 확인할 수 없는 상태다. 따라서 참고문헌[5]는 밸리-프리 라우팅 정책을 기반으로 발견적 해결법을 사용하여 AS 레벨 토폴로지에서 관찰되는 모든 사용자-제공자 링크를 완벽하게 찾아 내었다고 주장하고, 또한 놓칠 가능성이 있는 피어 링크 수를 추정할 수 있는 간단한 모델을 제공하였다. 본 논문에서는 IRL 연구실을 중심으로 참고문헌[5] 연구팀들이 제공하는 데이터 소스를 활용함을 밝힌다.

파워-로는 통신 네트워크에서 트래픽을 표현하기 위해 사용되어 왔으며, 네트워크 트래픽에서 나타나는 self-similarity와 중꼬리 분포 특징들이 파워-로와 관계한다. 만약 $P[X > x] = k^2 x^{-\alpha} L(x)$ (여기서, $k \in \mathbb{R}^+$ 이고 $L(x)$ 는 천천히 변화하는 함수 즉 $\lim_{t \rightarrow \infty} [L(tx)/L(x)] = 1$) 이면, 변수 X 는 중꼬리 분포를 따른다. 본 논문에서는 Faloutsos가 밝힌 AS 레벨 인터넷 토폴로지 상에서 세가지 파워-로를 대상으로 국내 AS 레벨 토폴로지 특성을 확인한다. 이들 세가지 파워-로 및 관련 정의를 간략히 기술하면 다음과 같다[1].

파워-로1 (순위지수 (rank exponent)) 노드 v 의 차수 d_v 는 노드 순위 r_v 의 상수 α 의 거듭제곱에 비례한다, $d_v \propto r_v^{-\alpha}$.

정의1: 그래프 노드들을 그들 차수를 기준으로 내림차순으로 정렬한다. 그리고 로그-로그 스케일에서 이들 노드의 순위에 대한 노드 차수 그림의 기울기를 순위지수 α 로 정의한다.

파워-로2 (차수지수 (degree exponent)): AS 노드 차수 d 의 빈도수 f_d 는 차수의 상수 β 의 거듭제곱에 비례한다, $f_d \propto d^{-\beta}$.

정의2: 로그-로그 스케일에서 차수에 대한 차수 빈도수 그림의 기울기를 차수지수 α 로 정의한다.

파워-로3 (고유치지수 (eigen exponent)): 그래프의 고유치 (eigenvalues) λ_i 는 순서 i 의 상수 ϵ 의 거듭제곱에 비례한다; $\lambda_i \propto i^\epsilon$.

정의3: 로그-로그 스케일에서 정렬된 고유치와 그들 순서에 대한 그림의 기울기를 고유치지수 ϵ 로 정의한다.

한편, 인터넷 노드 사이의 거리와 연결을 확인하기 위해 Faloutsos et al.는 홉 수에 따른 이웃 AS 노드의 개수를 조사하였다. 즉 h 홉 내 전체 AS 노드 쌍을 $P(h)$ 로 정의하고 자기 자신을 포함한 h 홉 내 포함된 전체 노드 쌍의 수를 나타내었다. 하나의 노드 쌍은 각 노드에서 한번씩 고려되기 때문에 두 번 계산된다. 따라서 $P(0) = N$ 와 $P(\delta) = N^2$ 의 관계가 성립하며, 여기서 N 은 전체 AS 노드 수이고, δ 는 AS 토폴로지 그래프의 직경이다. 이와 관련하여 다음과 같은 홉지수(hop-plot exponent)를 제시하였다.

근사1(홉지수): h 홉 내 전체 노드 쌍의 수 $P(h)$ 는 홉 수의 상수 \mathcal{H} 의 거듭제곱에 비례한다; $P(h) \propto h^{\mathcal{H}}, h \ll \delta$.

정의4: 로그-로그 스케일에서 홉 수에 대한 h 홉 내 전체 노드 쌍을 $P(h)$ 를 그린다. $h \ll \delta$ 에 대해 이 그림의 기울기를 홉지수 \mathcal{H} 로 정의한다.

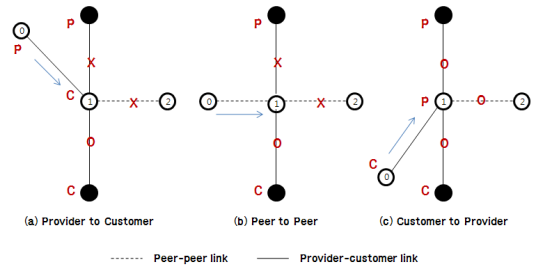
정의5: N 개 노드, E 개 에지(edge), 그리고 \mathcal{H} 홉지수가 주어지면, 실질적인 직경(effective diameter) δ_{ef} 는 다음과 같이 주어진다.

$$\delta_{ef} = \left(\frac{N^2}{N + 2E} \right)^{1/\mathcal{H}}$$

2.2 밸리-프리 라우팅 정책

BGP 라우팅 결정은 대부분 라우팅 정책에 의

해 결정되며 가장 중요한 요소는 이웃 AS 사이의 사업적 관계이다. 이들 사업적 관계는 두 가지 주요 타입, 즉 사용자-제공자와 피어-피어(peer-peer)가 있다[3,7]. 사용자-제공자 관계는 사용자는 인터넷 사용을 위한 트래픽 송·수신을 위해 제공자에게 비용을 지불하고, 제공자 AS는 사용자 AS에게 모든 인터넷 경로를 알려주며 사용자 AS의 인터넷 트래픽을 전달하는 역할을 수행한다. 그러나 사용자 AS는 자신의 제공자 AS 트래픽을 전달하지는 않는다. 즉 (그림 1)(a)에서 보듯이 자신의 제공자 트래픽은 자신의 하위 사용자 AS로만 전달하며 또 다른 제공자 AS 나 피어 AS로 전달하지 않는다.



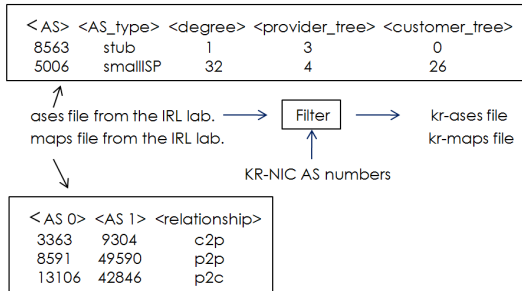
(그림 1) AS 레벨 ‘밸리-프리’ 라우팅 정책 (P: 제공자, C: 사용자, O: 허락, X: 거절)

한편, 피어-피어 관계는 일반적으로 “settlement-free”라고 표현하는데, 두 피어 AS에 직접 연결된 하위 스트림 AS에서 시작되고 종료되는 트래픽을 상호 간에 비용을 지불하지 않고 교환한다. 따라서 (그림 1)(b)에서 보듯이 AS는 피어-피어 링크를 포함하는 루트를 자신의 제공자 혹은 다른 피어 AS에게 전달하지 않는다. 이러한 특징을 요약하면, 오늘날 인터넷에서 사용하는 기본 정책은 “밸리-프리” 혹은 “노-밸리(No-valley)” 라우팅 정책이다. 즉 AS는 자신의 제공자 AS의 트래픽을 자신의 하위 사용자 AS로 전달하는 것 이외의 트래픽 전달 서비스는 허용하지 않는다. 마지막으로 사용자 트래픽을 받는 제공자 AS는 (그림 1)(c)와 같이 모든 인터넷 경로로 사용자 트래픽을 전달할 수 있어야 한다.

3. 국내 AS 레벨 인터넷 토폴로지

3.1 AS 레벨 토폴로지 추출 방법

본 논문에서는 최근 국내 인터넷 AS 레벨 토폴로지 특징을 분석하기 위해 미국 UCLA IRL 연구실이 제공하는 2009년 12월 2일자 실제 인터넷 AS 토폴로지 데이터 소스[11]를 사용한다. 전 세계적 인터넷 토폴로지 데이터 소스 내에서 국내 AS 레벨 토폴로지만 추출하기 위해 KRNIC이 제공하는 국내 AS 번호 리스트를[14] 기준으로 IRL에서 제공하는 AS 노드 파일 및 맵 파일을 재구성하였다(그림2) 참조). 따라서 국내 AS 파일(kr-ases)로부터 노드유형, 차수, 제공자트리, 그리고 사용자트리 정보를 각각 확인할 수 있다. 한편 국내 맵파일(kr-maps)로부터 AS 노드들의 상호연결 정보 및 링크 종류(사용자-제공자(c2p), 제공자-사용자(p2c), 그리고 피어-피어(p2p))를 확인할 수 있다.



(그림 2) 국내 AS 노드 및 맵 파일 획득

AS 레벨 토폴로지는 계층적 구조를 가진다. 즉 상위계층에 위치하는 AS 노드는 사용자 링크를 통해 하위 AS 노드들의 트래픽을 전달하는 서비스를 제공하게 된다. 과거에는 각 AS 노드의 차수 혹은 AS 프리픽스(prefix) 수를 기준으로 AS 계층을 분류하기도 하였으나, 최근에는 제공자-사용자의 관계를 추론하여 계층화하는 방법을 사용하고 있다[5,6]. 본 논문에서도 Oliveira et al.[5,6]가 제안한 방법을 사용하여, 각 AS의 하위 스트림 사용

자 AS의 수를 기준으로 다음 <표 1>과 같이 4계층으로 각 AS를 분류하였다. <표 1>에서 전체 AS 노드들의 집합 S 의 크기는 659이며, 세 번째 열의 $C(j)$ 는 AS 번호 $j \in S$ 의 하위 스트림 사용자 AS 개수를 나타낸다. <표 1>로부터 국내 721개 AS 중 62개 AS가 IRL 데이터 소스에서 확인되지 않고 있다. 이러한 이유는 IRL의 모니터링 지점들이 완벽하게 AS 토폴로지를 확인하지 못하고 있기 때문이거나, AS 번호를 할당 받고 있지만 현재 운영하지 않는 AS 이거나, 혹은 상위 제공자와 포인터-포인터로 연결하고 BGP를 운영하지 않는 AS 사이트에 해당한다. 실질적으로 망 제공자들이 사용자에게 AS 번호를 할당해주고도 정상적인 운영 서비스를 제공하지 않는 경우를 직접 확인할 수 있었다. 본 논문에서는 IRL 데이터 소스에서 확인된 659개 국내 AS 노드를 기준으로 파워-로 특징들을 확인하였다. 한편, largeISP에 속하는 5개 AS는 국내 주요 IX (Internet eXchange) 기관 - KT-IX (한국통신), Dacom-IX (데이콤) - 과 에스케이 텔레콤을 포함한 대형 ISP들이다.

(표 1) 2009년 12월 2일자 AS 레벨 토폴로지 데이터소스

ASE타입	AS 수(퍼센트)	기준
Tier1	0 (0)	제공자가 없는 AS
Large ISP	5 (0.8)	$C(j) \geq 50$
Small ISP	12 (1.8)	$5 \leq C(j) < 50$
Stub	642 (97.4)	$C(j) < 5$

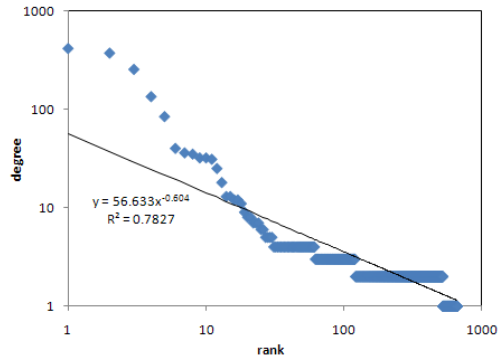
한편, 마이크로소프트(사)의 개방형 데이터베이스 연결 (ODBC: Open Database Connectivity)을 이용한 자체 제작한 MS Access DB 프로그램을 통해 앞에서 얻어진 국내 토폴로지 데이터 셋을 중심으로 AS 레벨 라우팅 패스를 다음 (그림3)과 같이 확보하였다. 그림에서 보는 바와 같이 하나의 레코드를 통해 두 AS 노드 쌍의 홉 수 (Metric 필드)와 라우팅 패스 (ASPath 필드)를 각각 확인할 수 있다.

FromASID	ToASID	Metric	Mask	BroadcastFlag	Source	ASPath	PerMetric
4008	11027	2	yes	peer	4008112811182		1
4008	11028	4	yes	provider	40081127917290311183		1
4008	11028	2	yes	provider	40082199111184		0
4008	11028	4	yes	provider	40081127917290311185		0
4008	11028	4	yes	provider	40081127917290311186		1
4008	11027	3	yes	peer	400821991104611187		1
4008	11028	3	yes	provider	40081127917290311188		1
4008	11028	4	yes	provider	40081127917290311189		1
4008	11028	3	yes	provider	40081127917290311190		1
4008	11028	3	yes	provider	40081127917290311191		1
4008	11027	4	yes	provider	40081127917290311192		1

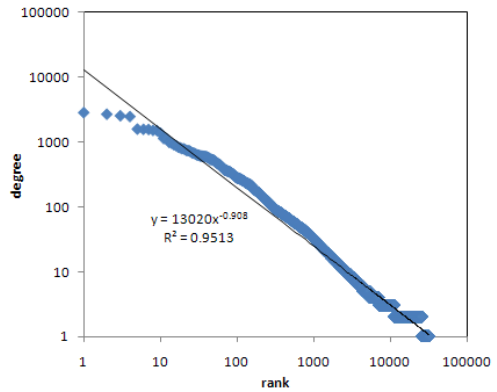
(그림 3) AS 레벨 라우팅 패스

3.2 파워-로 특성

본 절에서는 2.1절에서 설명한 세가지 파워-로 특성을 국내 AS 레벨 인터넷 토폴로지에 적용해 분석한 결과를 소개한다. 먼저, (그림 4)의 다이아몬드 형 점들은 AS 노드 차수를 내림차순으로 정렬하여 각 AS v ($v \in S$)에 순위 r_v 를 부여하고, 로그-로그 스케일로 순위에 따른 차수의 변화를 보여준다. 한편 (그림 4)의 실선은 최소제곱법 (least-squares)으로 근사(fitting)한 결과를 보여준다. 그러나 참고문헌[1]의 결과와 다르게 (그림 4)에서 보는 바와 같이 중꼬리 분포이긴 하지만 상관계수 0.884로 선형회귀선으로 예측하기 쉽다. 이러한 결과는 Chen이 지적한 결과와 매우 유사하다. 그러나 (그림 5)에서 보듯이 전 세계에 분포한 AS를 대상으로 조사하면 상관계수 0.975로서 참고문헌[1]에서 제시한 0.974와 비교해 거의 변화가 없음을 확인할 수 있다. 따라서 1998년과 비교해 AS 노드 수가 10배 이상 증가한 2009년12월2일 현 시점에도 Faloutsos의 AS의 차수 순위에 따른 파워-로는 여전히 유효함을 확인할 수 있다. 결론적으로, (그림 4)로부터 우리는 하나의 지협적인 AS 집합을 대상으로 할 경우, AS의 차수 순위에 따른 파워-로 특성이 매우 확연하게 나타난다고 볼 수는 없다(적어도 국내 AS 레벨 인터넷 토폴로지에서는).



(그림 4) 내림차순으로 정렬된 AS 노드의 차수 분포 (국내)

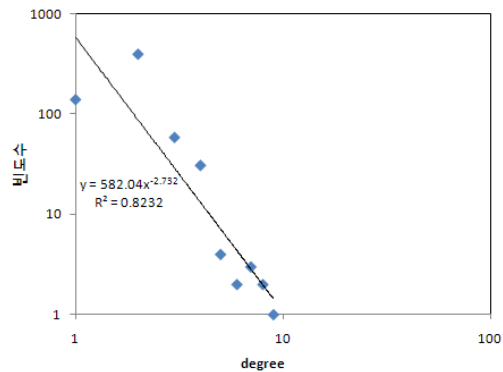
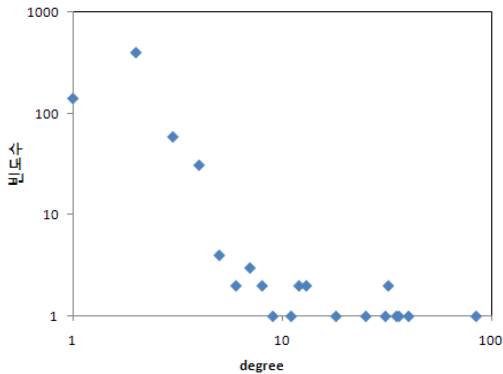


(그림 5) 내림차순으로 정렬된 AS 노드의 차수 분포 (전세계)

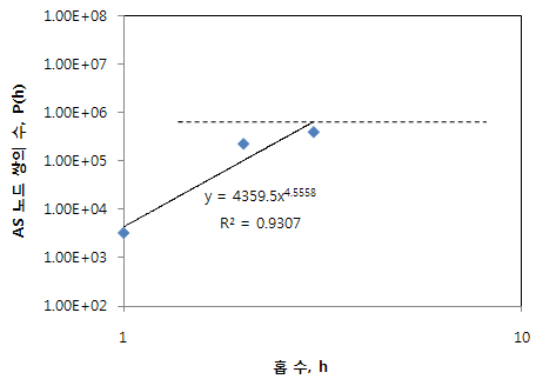
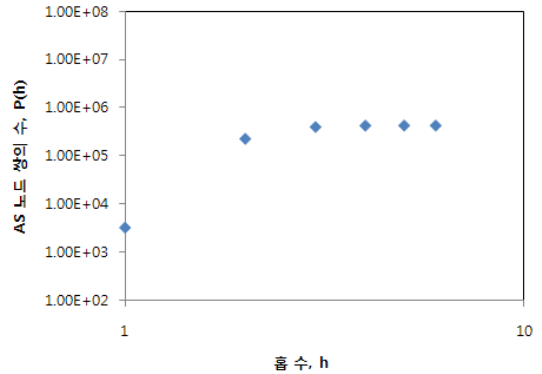
(그림 6)은 로그-로그 스케일에서 다이아몬드 형 점들은 차수 d 에 따른 빈도수 f_d 를 나타내었다. (그림 6)의 위쪽 그림은 659개 AS를 모두 나타내었으며 아래쪽 그림은 이들 중 빈도수가 아주 적은 높은 차수를 가진 16개를 제외한 97.6%인 643개 AS를 나타내었다. (그림 6)의 아래쪽 그림에서 실선은 최소제곱법으로 근사한 것이며, 그림에서 보듯이 상관계수가 0.907로 비교적 실측 값을 잘 나타낸다. 한편, 차수지수 α 는 -2.7로 나타났다. 따라서 국내 인터넷 토폴로지 역시, AS의 차수가 아무렇게나 분포하는 것이 아니라 낮은 차수를 가진 AS가 상당히 많으며 2.1절의 파워-로 2 특성을 이용해 이러한 분포를 비교적 정확하게 예측 가능하다.

인터넷 노드 사이의 거리와 연결을 확인하기 위해 Faloutsos et al. [1]는 홉 수에 따른 이웃 AS 노드의 개수를 조사하였다. 국내 환경을 적용하면, $P(0) = 659$ 와 $P(6) = 434,281$ 가 된다. (그림 7)은 로그-로그 스케일에서 다이아몬드 형 점은 홉 수 h 에 따른 노드 쌍의 수 $P(h)$ 를 보여준다. (그림 7)의 위쪽 그림은 가능한 전체 홉 수를 모두 나타내었으며 최대 홉 수는 '6'이 된다. 따라서 국내 AS 노드들은 6개 홉 내에서는 모두 접근 가능한 연결 구조를 유지하고 있다. 한편, (그림 7)의 아래쪽 그림은 $h < 4$ 에서 최소제곱법으로 근사 시켜 실선으로 나타내었으며 상관계수가 0.965으로 나타났다. 한편, 홉지수 $\beta = 4.5$ 로써 참고문헌[1]의 결과와 거의 일치되고 있으며, 따라서 국내 인터넷 연결구조가 약 4,000개 노드 수준의 1900년 말 세계 인터넷 연결구조와 거의 흡사하다고 볼 수 있다.

제2.1절 $P(h)$ 의 정의에 따라, $h = 1$ 에서 자신의 노드와 그리고 각 그래프의 에지가 두 번 고려된다. 따라서 $P(1) = N + 2E$ 의 관계식이 성립하며 (그림 7)에 나타난 최소제곱법으로 근사한 초기값 4359.5가 얼마나 정확한지 직접 확인해 볼 수 있다. 본 논문에서 고려되고 있는 국내 AS 레벨 인터넷 토폴로지 구성은 $N = 659$ 그리고 $E = 1304$ 이며, 따라서 최소제곱법 근사 시 초기치 오차가 33% 수준이다. 이러한 심각한 초기치 오차가 발생하는 이유는 본 논문에서 사용하고 있는 UCLA IRL 연구실 제공 데이터 소스가 아직도 상당한 수의 링크들을 제대로 파악하지 못하고 있는 결과다. 이와 같은 문제점은 참고문헌[1]과 [2]에서도 언급하고 있으며, 특히 stub AS 노드의 경우 상당수의 피어 링크들이 정확하게 파악되지 못하고 있다.



(그림 6) AS 노드 차수 d 에 따른 빈도수 f_d
(위: 전체, 아래: 전체 AS 중 97.6%)



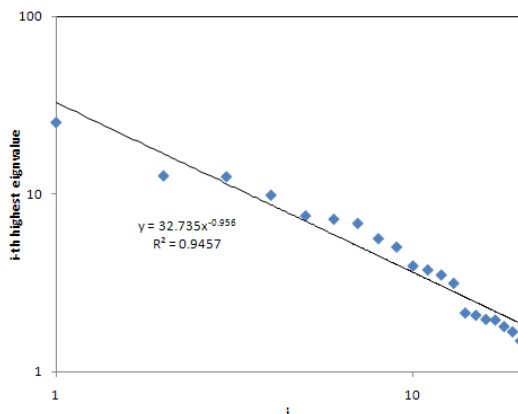
(그림 7) 홉 수 h 에 따른 AS 노드 쌍의 수 $P(h)$
(위: 전체, 아래: $h < 4$)

한편 제2.1절 (정의5)에 설명한 그래프의 실질적인 직경 $\delta_{s,r}$ 는 AS 레벨 토폴로지에서 하나의 AS 노드로부터 “충분히 많은” 다른 AS 노드로 전파되는데 소요되는 홉 수를 의미한다. 즉 두 AS 노드가 $\delta_{s,r}$ 홉 내에 존재할 확률이 매우 높다는 것을 의미한다. 국내의 경우, (정의5)에 의해 다음과 같이 계산된다.

$$\delta_{s,r} = \left(\frac{659^2}{659 + 2 \cdot 1304} \right)^{1/4.5558} = 2.92$$

따라서 실질적인 직경 값의 실링(ceiling)을 취한 홉 수 3을 기준으로 볼 때, 3 홉 내에 존재하는 AS 노드 쌍들은 전체 91% 이상이 된다.

그래프의 고유치는 그래프의 인접행렬(adjacency matrix)에 대한 고유치이며, 기본적인 그래프 특징들을 잘 나타낸다. 즉 그래프의 직경, 예지 수, 스패닝 트리 수, 연결된 컴포넌트, 그리고 노드 사이의 거리 등을 그래프의 고유치를 통해 확인할 수 있다[15]. (그림 8)은 국내 AS 레벨 인터넷 토폴로지를 대상으로 고유치를 구해 내림차순으로 정렬한 후 상위 20개 고유치를 나타내었다(다이아몬드 형 점). 한편, 실선은 최소제곱법으로 근사한 그림이며 상관계수가 0.97로 제2.1절에서 설명한 파워-로3을 정확하게 확인할 수 있다.



(그림 8) 내림차순으로 정렬된 고유치(eigenvalue)

4. 결론 및 토의

본 논문에서는 국내 AS 레벨 인터넷 토폴로지의 파워-로 특성을 조사하였다. 이를 위해 UCLA IRL 연구실이 제공하는 2009년 12월 2일 실제 AS 레벨 데이터 소스를 이용해 국내 AS 번호(KRNIC 제공)를 대상으로 AS 노드와 링크 정보를 추출했다. 국내 AS 레벨 인터넷 토폴로지 역시 파워-로를 이용해 AS 노드 차수, 홉 수에 따른 노드 쌍의 수, 그리고 그래프의 고유치 분포를 비교적 정확하게 나타낼 수 있음을 확인하였다. 특히, 이들 파워-로 근사는 상관계수가 각각 90.7%, 96.5%, 그리고 97%로 조사되었다.

AS 레벨 인터넷 토폴로지 연구는 기존의 IP 혹은 라우터 레벨 인터넷 토폴로지 연구와 비교해 그 규모가 훨씬 광범위하다. 즉 전역적 네트워크(network-wide) 규모에서 트래픽 흐름이나 혹은 통신 프로토콜 성능을 평가할 때, AS 레벨 인터넷 토폴로지가 라우터 레벨 인터넷 토폴로지와는 차별화된 장점들을 제공한다. 국내의 경우, 아직 KR-AS 레벨 인터넷 토폴로지 형상 조차 제대로 파악된 연구 결과를 찾아보기가 불가능한 실정이었다. 본 논문은 비록 UCLA IRL 연구실이 제공하는 데이터 소스를 사용했지만, 국내 AS 레벨 인터넷 토폴로지의 기본적인 특성을 최초로 밝혔다는 점에서 의미가 있다. 따라서 국내 AS 레벨 토폴로지 모델이 필요한 연구 분야에서는 본 논문에서 제시한 세가지 파워-로 특성과 관련된 파라미터가 매우 중요한 의미를 가질 것이다. 그러나 제3.2절에서 설명한 바와 같이 보다 정확한 국내 AS 토폴로지 확보를 위해 주요 지점의 라우터와 BGP 세션을 형성해 정확한 자료 수집이 절대적으로 필요한 시점이다. 마지막으로 본 논문이 국내 연구진들이 AS 레벨 인터넷 토폴로지에 관심을 가질 수 있는 출발점이 되기를 희망한다.

참 고 문 헌

- [1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On Power-Law Relationships of the Internet Topology," In Proc. of ACM SIGCOMM, 1999, pp. 251-262.
- [2] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, "Power-Laws and the AS-level Internet Topology," IEEE/ACM Trans. On Networking, Vol.11, No.4, 2003, pp.514-524.
- [3] L. Gao, "On Inferring Autonomous System Relationships in the Internet," IEEE/ACM Transactions on Networking, Vol.9, Issue 6, 2001, pp.733-745.
- [4] I.V. Beijnum, "Building Reliable Networks with the Border Gateway Protocol", O'Reilly, 2002.
- [5] R. Oliveira, D. Pei, W. Willinger, B. Zhang and L. Zhang, "In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure," In Proc. of SIGMETRICS'08, 2008.
- [6] R. Oliveira, D. Pei, W. Willinger, B. Zhang and L. Zhang, "Quantifying the Completeness of the Observed Internet AS-level Structure," Technical Report TR-080026, Computer Science Department UCLA, 2008.
- [7] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet AS-level topology," SIGCOMM Computer Communication Review, Vol.35, Issue 1, 2005, pp. 53-61.
- [8] Q. Chen, H. Chang, R. Govindan, and S. Jamin, "The origin power laws in Internet topologies revisited," In Proc. of INFOCOM 2002, 2002
- [9] RIPE routing information service project, available at <http://www.ripe.net>
- [10] Route Views routing table archive, available at <http://routeviews.org>.
- [11] UCLA IRL Internet Topology Collection, available at <http://irl.cs.ucla.edu/topology>
- [12] J. Winick, S. Jamin, "Inet-3.0: Internet Topology Generator. Technical Report, Department of EECS, University of Michigan, 2002.
- [13] C.D. Murta, J.N. Maciel, NIT: The New Internet Topology Generator. Technical Report, Department of Informatics, University Federal of Parana, 2008.
- [14] 국내 AS 번호 목록, available at <http://www.ip.kisa.or.kr/main.html>.
- [15] . M. Cvetkovic, M. Boob, and H. Sachs, "Spectra of Graphs," Academic press, 1979.

● 저 자 소 개 ●



강 구 흥

1985년 경북대학교 전자공학과 졸업(학사)
 1990년 충남대학교 대학원 전자공학과 졸업(석사)
 1998년 포항공과대학교 대학원 전자계산학과 졸업(박사)
 1985년 ~ 1999년 한국전자통신연구원(ETRI) 선임연구원
 2000 ~ 현재 서원대학교 정보통신공학과 교수
 관심분야 : 성능평가, 컴퓨터 네트워크, 네트워크 보안
 E-mail : khkang@seowon.ac.kr