

## Topologia e Modelagem Relacional da Internet Brasileira

Nilton Alves<sup>1,2</sup>, Márcio P. de Albuquerque<sup>1</sup>, Marcelo Portes de Albuquerque<sup>1</sup>  
e Joaquim Teixeira de Assis<sup>2</sup>.

Centro Brasileiro de Pesquisas Físicas - CAT-CBPF/MCT  
Coordenação de Engenharia e Operações da Rede-Rio de Computadores  
Rua Dr. Xavier Sigaud n.150 - Urca - 22290-180 - Rio de Janeiro, RJ

Instituto Politécnico da Univ. do Estado do Rio de Janeiro - IPRJ/UERJ  
Rua Alberto Rangel, s/n, Vila Nova - 28630-050 - Nova Friburgo, RJ

mpa, naj, marcelo@cbpf.br, joaquim@iprj.uerj.br

**Abstract.** *In this paper we present a local vision of the Brazilian Internet topology and its relations with the world Internet, in order to characterize regional statistical properties of the network. The description of this topology is based on the self-organized network formed by Brazilian Autonomous Systems, obtained from BGP inter domain routing tables in Rio Metropolitan Network routers. We present the properties of the local organization and its relationship with the scale-free network model proposed by Barabási. The power-law coefficients from rank and neighbors distributions and the Brazilian network map are presented and discussed.*

**Resumo.** *Neste trabalho abordamos uma visão da topologia da Internet brasileira e sua relação com a Internet mundial, com o objetivo de modelar as propriedades topológicas regionais e seus comportamentos estatísticos. As análises são realizadas na rede auto-organizada formada pelo relacionamento entre os sistemas autônomos brasileiros em dados obtidos a partir da tabela BGP completa em roteadores da Rede-Rio. Apresentamos as propriedades da organização local e seu relacionamento com o modelo em escala livre proposto recentemente por Barabási. As medidas dos coeficientes das leis em escala-livre para a organização em ranque, a distribuição do número de vizinhos e o mapa da topologia da Internet brasileira são apresentados e discutidos.*

# 1. Introdução

Diversas organizações em redes estão presentes em nosso cotidiano, e.g., as redes de distribuição elétrica, redes sociais, redes rodoviárias, redes de computadores, redes de neurônios, etc. Até recentemente os estudos em sistemas em redes eram principalmente simulados por uma área da matemática chamada teoria dos grafos. A falta de dados experimentais de tais sistemas levavam a uma abordagem por meio de uma estrutura aleatória, conhecida por *random-graphs*. As primeiras análises de estruturas em redes foram introduzidas por P. Erdős e A.Rényi nos anos 60. O modelo proposto na época consistia de  $N$  nós, interconectados entre si com uma probabilidade  $p$ . Através desse tipo de consideração uma rede aleatória segue uma distribuição de Poisson, fazendo com que seja raro encontrar nós com uma grande concentração de conexões. O trabalho de Albert-László Barabási, introduziu um outro tipo de estrutura em rede, as redes em escala-livre, *scale-free networks* [1]. O estudo de Barabási mediu a distribuição de nós em páginas da Internet e mostrou que diversas redes têm topologias com conexões preferenciais. Em outro trabalho, Faloutsos et al [2] apresentaram um estudo da topologia da Internet mundial e sua distribuição em lei de potência. As observações eram feitas entre Sistemas Autônomos (ASs) e roteadores, por meio de dados coletados de roteadores públicos mantidos pelo Laboratório Americano para Pesquisas e Aplicações em Redes de Computadores (NLANR) [3]. Na mesma linha de estudo, Pastor-Satorras [4] apresentou as propriedades dinâmicas e as correlações de vizinhança da Internet. Os resultados obtidos apresentam um comportamento que também segue uma lei de potência, com coeficiente igual a  $-0,82$  para a análise ordenada do número de conexões e  $-2,1$  para a distribuição de conexões de ASs. Os estudos apresentados por Barabási e Pastor-Satorras consideram a Internet como um processo complexo e dinâmico resultado da topologia observada.

As motivações que nos levaram a este trabalho foram caracterizar as propriedades da topologia da Internet e o relacionamento entre ASs no Brasil. Seria a estrutura local equivalente a da Internet mundial? Qual o tamanho dos *backbones* da Internet hoje no Brasil e as suas relações com a Internet Mundial? Seriam estas propriedades estacionárias? As propriedades observadas em escala regional se repetem em larga escala (fractalidade)?

Este artigo está organizado da seguinte maneira: a seção 2 apresenta aspectos fundamentais dos sistemas complexos baseados em leis de potência e suas observações na Internet mundial; a seção 3 descreve a metodologia para a obtenção dos dados através da tabela completa (*full-routing*) entre sistemas autônomos; a seção 4 apresenta os resultados para as redes brasileiras bem como uma análise comparativa em relação à Internet mundial e por fim a seção 5 é dedicada às conclusões e discussões finais do trabalho.

## 2. Sistemas Complexos

A área da física dedicada a mecânica estatística vem oferecendo novos modelos matemáticos que permitem estudar e descrever a estrutura e a dinâmica de vários parâmetros da Internet. Sistemas complexos se caracterizam por comportamentos matemáticos que podem ser descritos por leis de potência. Atualmente vários grupos da ciência da computação vêm desenvolvendo modelos de redes com diferentes topologias, veja por exemplo o trabalho de Medina et al. [5] e as referências apresentadas aos modelos Waxman e Transit-Stub. O modelo em escala-livre proposto por Barabási propõe dois mecanismos que justificam o aparecimento de leis de potência na topologia de redes de computadores: crescimento incremental da rede e a existência de pontos de conexões preferenciais. O crescimento incremental se refere a uma topologia em rede aberta, a qual é formada pela anexação contínua de novos nós. A conexão preferencial se refere a tendência apresentada por novos nós em se conectarem aos nós com mais conexões, isto é, os “mais populares” da rede.

No entanto, existem diferenças fundamentais entre redes com distribuição aleatória e redes em escala-livre. Nas redes aleatórias a maior parte dos nós tem aproximadamente o mesmo número de conexões,  $k \sim \langle k \rangle$ . No caso de redes em escala-livre muitos nós terão poucas conexões enquanto somente alguns nós terão muitas conexões ganhando importância no cenário observado. Estas considerações são importantes pois, as medidas baseadas em valores mínimos, médios e máximos não são ideais para caracterizar este segundo tipo de topologia. Estes valores descartam informações relevantes que estão diretamente associados a distribuição em questão. Um parâmetro de relevância que descreve esta distribuição é o coeficiente da lei de potência.

Uma distribuição em lei de potência tem a forma  $y \propto x^B$  e permite uma caracterização compacta da topologia de rede. Se a relação  $(x, y)$  for apresentada em um gráfico em escala  $\log - \log$ , o expoente da lei de potência é definido pela inclinação da reta observada no gráfico. Este pode ser considerado como a “assinatura” característica de um fenômeno em lei de potência.

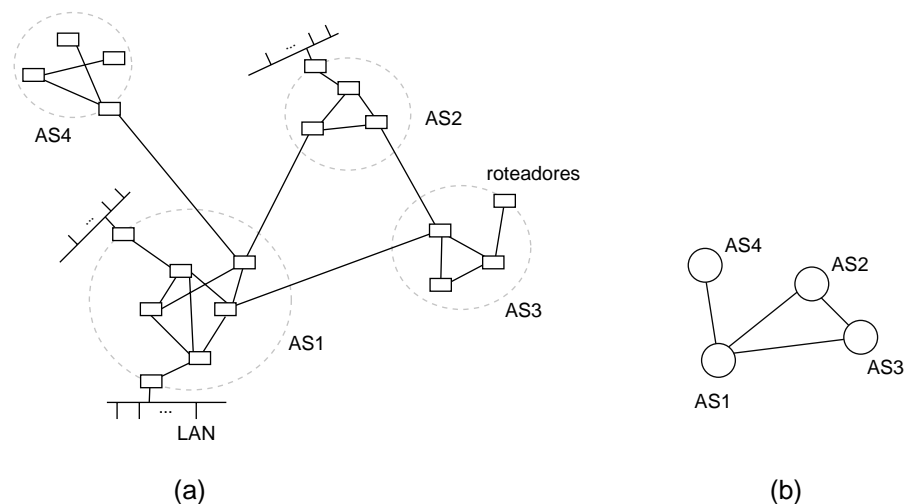
## 3. Análise BGP

O funcionamento da Internet depende do uso do protocolo BGP\* [6]. Este protocolo estabelece mecanismos de anúncios de prefixos de rede entre ASs. Os relacionamentos acontecem após o estabelecimento de um acordo de troca de tráfego com ou sem compromisso financeiro (*peering*) entre dois ou mais ASs. Em qualquer modelo de *peering* usado, é preciso que cada AS determine quais blocos de redes vai anunciar para os vizinhos BGP, estabelecendo assim um caminho entre ASs para que todos os prefixos de redes (LANs) sejam alcançados por todos os usuários da rede. Um AS, através de sua

---

\*BGP - Border Gateway Protocol.

política de roteamento, define também quais prefixos vai receber dos anúncios dos ASs vizinhos, Figura 1.



**Figura 1: Estrutura da Internet - (a) destaque para a estrutura interna dos Sistemas Autônomos (ASs) e (b) destaque para a estrutura de relação entre os ASs.**

Neste trabalho utilizamos as informações do protocolo de roteamento BGP obtidas no roteador de borda da Rede-Rio que dispõe da tabela completa de prefixos de rede da Internet. Dados equivalentes podem também ser obtidos em roteadores (*route servers*) de acesso livre disponibilizados nos principais ASs. Um pré-processamento na tabela BGP é realizado a fim de retirar a redundância de rotas e as políticas de custo de um determinado caminho (*prepend*) que alteram a profundidade da rede. A partir dessa tabela excluímos todos os ASs exceto aqueles que correspondem a provedores Internet brasileiros, i.e., aqueles cadastrados no registro de endereçamento da América Latina (LACNIC)<sup>†</sup>. Podem existir ainda alguns ASs estrangeiros em operação no Brasil, porém estes não foram considerados neste estudo pois não estão oficialmente cadastrados na base de registros de endereços do LACNIC.

#### 4. Resultados das observações

A partir da análise da tabela BGP concluímos que existem na Internet (Fevereiro de 2004) 16.538 ASs em funcionamento, sendo que 164 (0,99%) são registrados por *backbones* brasileiros. Destes ASs brasileiros, 26 (16%) têm conexões com ASs fora do Brasil. O número total de conexões de ASs brasileiros é de 340, sendo 99 (29%) com ASs fora do Brasil. O índice de conexões de ASs brasileiros entre si é de 1,80% (241), em um total de  $N(N - 1)/2 = 13.366$  possíveis conexões. Na Internet mundial este mesmo índice é de 0,021% (27.925) em um total de 136.744.453 possíveis conexões.

<sup>†</sup>LACNIC - Registro de Endereçamento de Internet para América Latina e Caribe.



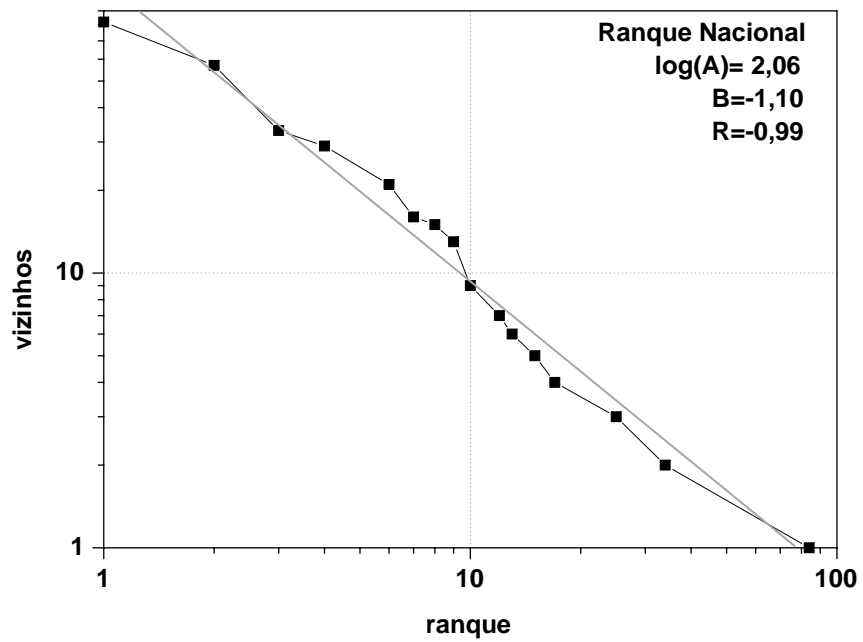
fractais.

Os gráficos da Figura 4 apresentam as distribuições dos números de vizinhos  $k$  de ASs (brasileiro e internacional), isto é, a probabilidade  $P(k)$  de ASs com o mesmo número de vizinhos. Na construção do gráfico da Figura 4(a) foram levados em conta ASs brasileiros com conectividade com outros ASs no Brasil e no exterior. Os ajustes em lei de potência podem ser observados nesta figura, com expoente  $B$  igual a  $-2,07$  e  $-2,08$  para as distribuições brasileira e internacional respectivamente. Este comportamento linear semelhante caracteriza novamente a fractalidade do sistema. Para o cálculo dos ajustes utilizamos 99,78% e 95,70% dos pontos nos gráfico 4(a) e 4(b) respectivamente, devido à baixa representatividade dos pontos de maior conectividade.

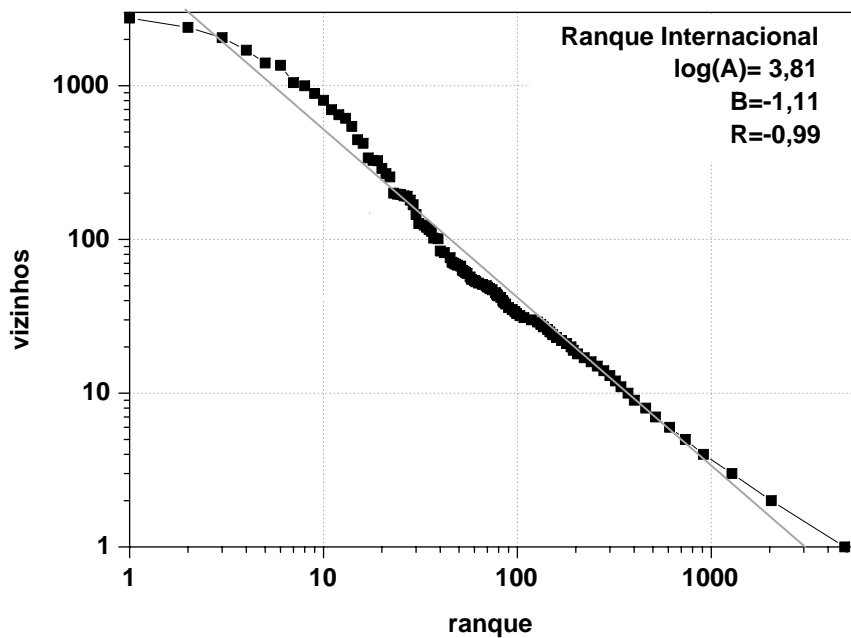
A partir da análise dos dados, podemos observar que uma característica das redes em escala-livre é a assimetria entre o número de ASs com muitos vizinhos em relação ao número de ASs com poucos vizinhos. Um parâmetro que pode estabelecer um ponto de separação entre essas duas classes é o valor médio de vizinhos dos ASs. O valor médio de conexões por AS obtido é igual a 3,54 e 3,37 para a Internet brasileira e mundial, respectivamente. No caso da Internet brasileira somente 14.6% dos ASs tem maior número de vizinhos que a média. No caso da Internet mundial este valor é igual a 7%. Esta diferença entre os valores em diferente escalas (regional e mundial) pode ser entendida devido ao crescimento do número de ASs com poucas conexões.

Os resultados encontrados apresentam as relações entre os ASs na rede brasileira e na Internet mundial. A organização em ranque enfatiza o tamanho dos ASs com muitas conexões (Figura 3) e a análise da distribuição estatística de vizinhos enfatiza os ASs com poucas conexões (Figura 4). Em ambos os casos e para as duas situações (nacional e internacional) o comportamento em escala-livre e em lei de potência pode ser caracterizado. O expoente  $B$  encontrado em todos os casos confirma que a Internet alcançou um estado estacionário comparativamente aos trabalhos de [2] e [4]. O comportamento regional segue o mesmo comportamento global caracterizando a Internet como uma rede independente da escala em que se observa.

Finalmente, neste trabalho mostramos que um processo que apresenta um ranque Zipf  $r$  do número de vizinhos  $v$ , tem da mesma forma uma função densidade de probabilidade em lei de potência. Os dados coletados mostram que no caso de um ranque com coeficiente  $B$  igual a  $-1$ , o expoente que exprime o comportamento da lei de potência da distribuição será igual a  $-2$ . A lei de Zipf ( $v \propto r^B$ ) está relacionada com a lei de Pareto ( $r \propto v^{1/B}$ ) pois são diferentes interpretações do mesmo fenômeno. A distribuição estatística representada por meio da lei de potência é obtida a partir da derivação da lei de Pareto. Este relacionamento explica os expoentes encontrados.



(a)



(b)

Figura 3: Gráficos em escala  $\log - \log$  do ranqueamento em ordem decrescente dos Sistemas Autônomos da: (a) Internet brasileira e (b) Internet mundial. A Internet brasileira apresenta o mesmo comportamento da internet mundial. Os dados correspondem a observações realizadas no início de 2004.

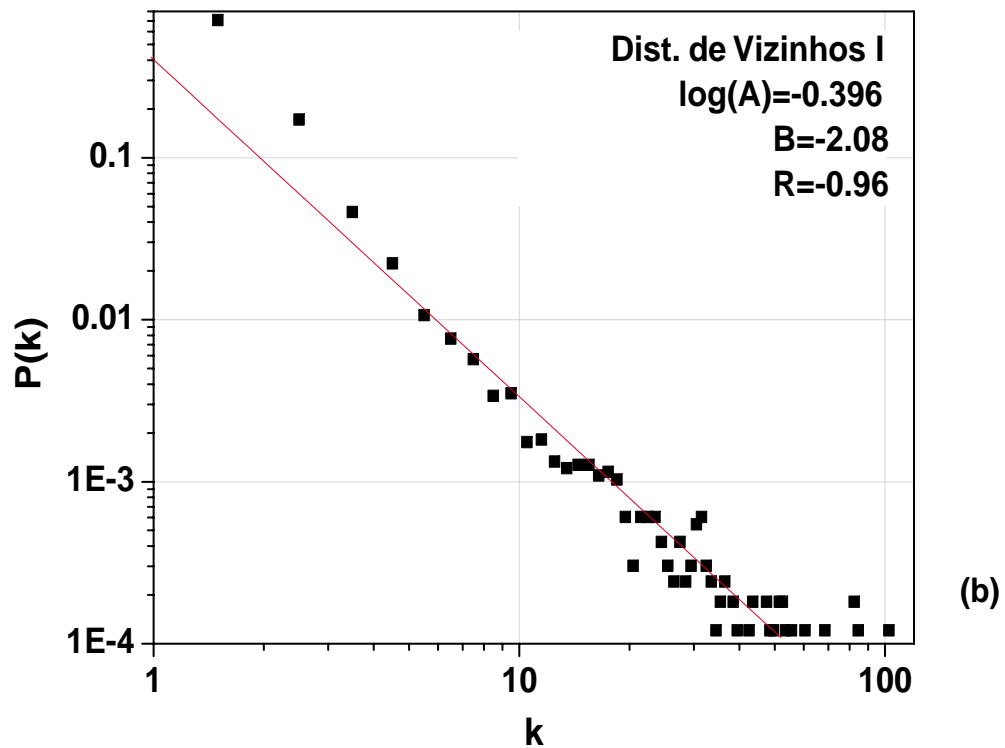
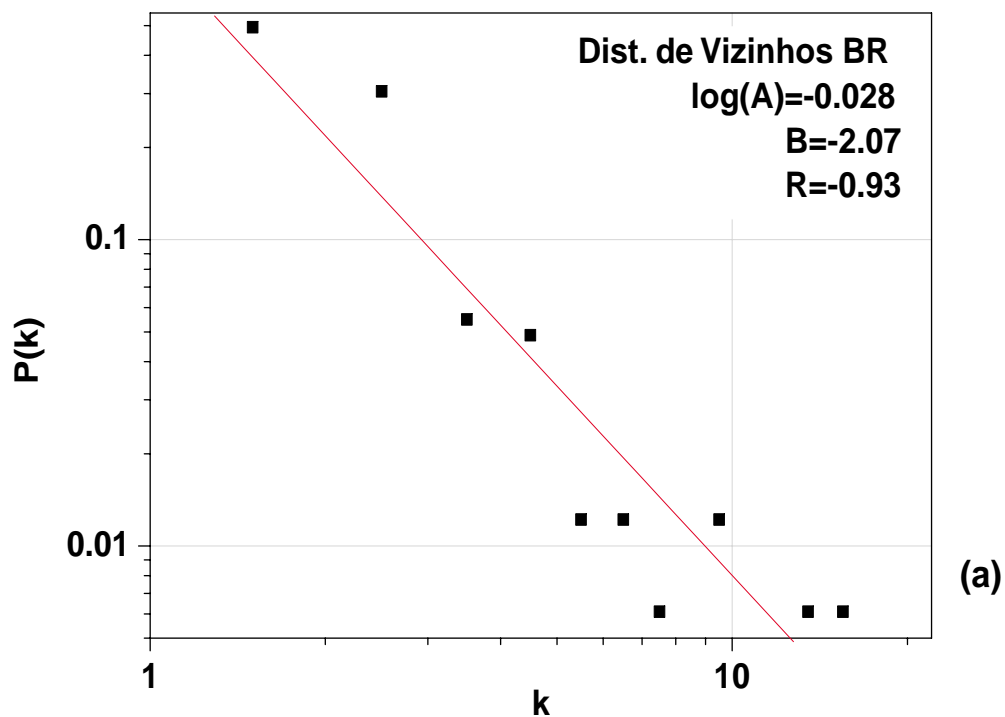


Figura 4: Gráfico em escala  $\log - \log$  da distribuição de ASs por número de vizinhos para a Internet mundial e brasileira. O ajuste fornece um coeficiente de  $-2,06$  e  $2,08$  para a internet brasileira e mundial respectivamente. As observações foram realizadas no início do ano de 2004.



## **5. Conclusões**

Neste trabalho apresentamos as características topológicas da Internet brasileira. Os dados foram processados através da tabela de roteamento completa entre sistemas autônomos, expressas pelo protocolo de anúncios de prefixos de redes BGP. A visão apresentada corresponde a uma análise a partir do AS da Rede-Rio de Computadores. Os resultados comprovam que a topologia da rede brasileira segue uma distribuição em escala-livre assim como a Internet mundial. A análise estatística desses resultados nos revela que a Internet Brasileira apresenta propriedades topológicas com distribuições em leis de potência.

É importante destacar que esta visão corresponde à observação a partir da Rede-Rio de Computadores. Sabemos que algumas poucas políticas de troca de tráfego acabam não aparecendo na tabela BGP observada neste ponto e esta tabela pode ser diferente de uma observação em outros pontos da rede. Uma visão geral da topologia brasileira poderia ser obtida através de uma análise complementar a partir de outros pontos. Uma análise das tabelas dos principais ASs poderia ser um bom começo.

O conhecimento da topologia de redes que seguem um comportamento em escala-livre pode por exemplo auxiliar os centros de operação e gerência de redes no combate à disseminação de vírus, no estabelecimento de políticas para prevenção de falhas, na determinação de pontos estratégicos para distribuição de vídeo ou no desenvolvimento de protocolos visando que os mesmos sejam eficientes na sua transmissão na própria rede.

Finalmente é possível dizer que o comportamento da topologia é auto-similar, i.e., é independente da escala em que ele é observado (fractal). Este trabalho comprova que o comportamento da topologia da rede brasileira, sua organização e acordos de troca de tráfego, seguem o mesmo comportamento estatístico da Internet mundial. As leis de potência conseguem descrever a Internet pois esta é uma grande rede cujo crescimento tem dependências diversas, de natureza tecnológica, econômica e política.

## **6. Agradecimentos**

Os autores gostariam de agradecer a Rede-Rio de Computadores da FAPERJ, o Ministério de Ciência e Tecnologia do Brasil e o registro de endereçamento de Internet para América Latina e Caribe (LACNIC).

## Referências

- [1] Albert-László Barabási; The Physics of the Web; Physics World, pg.33-38, Julho, 2001.
- [2] M. Faloutsos, P. Faloutsos e C. Faloutsos; On power-law relationships of the internet topology; ACM SIGCOMM, Computer Communication Rev., 29, pg.251-262, 1999.
- [3] The National Laboratory for Applied Network Research (NSF); Informações sobre tabelas de roteamento da Internet baseado em dados BGP. (veja <http://moat.nlanr.net>)
- [4] R. Pastor-Satorras, A. Vazquez, e A. Vespignani; Dynamical and correlation properties of the Internet; Physical Review Letter, vol. 87, pag.258701, 2001.
- [5] A. Medina, I.Matta e J. Byers; On the origin of power laws in the internet topologies; ACM SIGCOMM Computer Communication Rev., vol.30, n.2, pg.18-28, 2000.
- [6] Alex S. de Moura; Dicas na Configuração do Protocolo BGP-4 - Parte 1; Revista News-Generation - RNP, vol.5, n.1, Jan.2001.
- [7] L.A. Adamic e B.A. Huberman; Zipf's law and the Internet; Glottometrics 3, pg.143-150, 2002.