

Análise da Estabilidade de Ranqueamento de Grau na Rede de Sistemas Autônomos da Internet

Fernando F. Machado¹, Daniel R. Figueiredo², Miguel Elias M. Campista¹ *

¹GTA/PEE-COPPE/DEL-Poli - Universidade Federal do Rio de Janeiro (UFRJ)
Rio de Janeiro - RJ - Brasil

²LAND/PESC-COPPE - Universidade Federal do Rio de Janeiro (UFRJ)
Rio de Janeiro - RJ - Brasil

{ffraga,miguel}@gta.ufrj.br, daniel@land.ufrj.br

Abstract. *Many networks have properties that remain invariant over time, even when many vertices and edges are added and removed. One of such properties is the ranking stability of the most important vertices that can be obtained from network historical data. In this work, we analyze the degree ranking stability over time of the Internet Autonomous Systems network based on techniques from prior works. Additionally, we propose an empirical analysis to evaluate the influence of external factors to the network structure. Results show that the stability of top-ranked vertices is predominantly affected by external factors, while the remaining vertices are also affected by common degree variations in the network. Despite the Autonomous Systems network has been analyzed under several aspects, this work is the first to evaluate the temporal dynamics of the degree ranking.*

Resumo. *Muitas redes possuem propriedades que tendem a se manter ao longo do tempo, mesmo havendo grande variação na quantidade de vértices e arestas. Uma dessas propriedades é a estabilidade do ranqueamento dos principais vértices que pode ser avaliada através dos dados históricos da rede. Neste trabalho, técnicas anteriormente desenvolvidas são utilizadas como base para a análise da estabilidade temporal do ranqueamento de grau da rede de sistemas autônomos da Internet. Adicionalmente, uma análise empírica é proposta para avaliar a influência de fatores externos à estrutura dessa rede. Os resultados mostram que a estabilidade das posições do topo é predominantemente afetada pelos fatores externos, enquanto que as demais posições são também afetadas pelas variações de grau regulares da rede. Apesar da rede de sistemas autônomos ter sido analisada sob muitos aspectos, este trabalho é o primeiro a avaliar a dinâmica temporal do ranqueamento.*

1. Introdução

O estudo empírico de redes reais vem identificando propriedades que se repetem em redes de origens distintas. Uma das propriedades comuns a diversas redes reais, como redes tecnológicas e redes sociais, é a distribuição de grau com cauda pesada, na qual

*Este trabalho foi parcialmente financiado pela CAPES, CNPq, FAPERJ e FINEP.

há a ocorrência de vértices com grau muito superior à média [Figueiredo, 2011]. Desas redes, as que possuem distribuição de grau próxima a uma lei de potência, ou seja, seguindo o modelo $P(k) \sim k^{-\gamma}$ com $\gamma > 1$ constante, onde $P(k)$ denota a fração de vértices com grau k , são chamadas de redes livre de escala [Barabási e Albert, 1999]. Muitas redes apresentam essa propriedade, tais como a rede formada por *links* de páginas *web* [Albert et al., 1999], a rede de vizinhança de sistemas autônomos da Internet [Siganos et al., 2003] e até mesmo a rede formada por citações de artigos científicos [Redner, 1998].

Um importante aspecto de qualquer rede é o ranqueamento de seus vértices segundo algum critério de importância. Por exemplo, o PageRank é um algoritmo de ranqueamento inicialmente proposto pelo Google para ranquear os vértices da rede de páginas *web* em função da estrutura da rede [Page et al., 1999]. Entretanto redes reais estão constantemente se modificando com a adição e remoção de vértices e arestas. Surge assim o interesse na dinâmica do ranqueamento ao longo do tempo, com a evolução da rede. Em particular, o estudo da estabilidade do ranqueamento, que se refere a propriedade dos vértices manterem suas posições no ranqueamento inalteradas ao longo do tempo.

Uma característica estrutural que se mantém ao longo do tempo em muitas redes é a distribuição de grau de cauda pesada, e em particular na rede de sistemas autônomos que formam a Internet. Mas como essa característica influencia o ranqueamento por grau dos vértices da rede? Ghoshal e Barabási demonstram a existência de vértices com diferenciada estabilidade de ranqueamento, chamados de superestáveis, e ainda que é possível prever a quantidade aproximada desses vértices a partir de um grafo instantâneo da rede [Ghoshal e Barabási, 2011]. Já em outro trabalho, Blumm et al. provam que o nível de estabilidade geral do ranqueamento está relacionado ao ruído do sistema cuja amplitude absoluta pode ser estimada a partir do histórico das variações da característica analisada, que pode ser o grau [Blumm et al., 2012]. Dois aspectos não analisados nesses trabalhos são a variável tempo e o impacto de fatores externos à estrutura da rede, aqui chamados de pontuais. Tanto o tempo quanto os fatores pontuais, que podem afetar um único vértice da rede de maneira relevante, têm influência sobre o ranqueamento da rede mas não foram ainda modelados dada a complexidade.

Este trabalho propõe uma metodologia de análise da estabilidade do ranqueamento de grau considerando o impacto do tempo e dos fatores pontuais em redes de computadores, mais precisamente, na rede de sistemas autônomos da Internet. A metodologia é baseada em três etapas: (i) na análise da existência de vértices superestáveis considerando o tempo, (ii) na análise do impacto do ruído no nível de estabilidade do ranqueamento ao longo do tempo e (iii) na análise dos fatores pontuais que afetam o ranqueamento. Parte da metodologia é baseada nas técnicas dos trabalhos relacionados [Ghoshal e Barabási, 2011, Blumm et al., 2012] e parte é baseada na análise empírica dos dados da rede de sistemas autônomos da Internet. Os resultados confirmaram a presença de vértices superestáveis na rede de sistemas autônomos da Internet, além de demonstrar que os fatores pontuais afetam de forma importante a estabilidade do ranqueamento ao alterar a aptidão dos vértices, em particular quando se trata da rede de sistemas autônomos.

O restante deste trabalho está organizado da seguinte forma. A Seção 2 resume

os trabalhos relacionados. A Seção 3 detalha o conjunto de dados utilizado. A Seção 4 descreve a metodologia empregada na análise. A Seção 5 apresenta os resultados encontrados. A Seção 6 conclui este trabalho e apresenta as direções futuras.

2. Trabalhos Relacionados

O ranqueamento dos vértices de uma rede pode ser realizado em função de diferentes métricas de centralidade, que visam capturar a importância de um vértice na rede. Nessa direção, duas métricas bastante utilizadas são o grau e o PageRank [Boldi e Vigna, 2014]. Apesar do grau inferir a centralidade baseado no número de adjacências de um dado vértice, ele é frequentemente utilizado pela simplicidade e por exigir apenas conhecimento local. O PageRank foi proposto pelo Google com o objetivo de calcular a importância de páginas *web* a partir da estrutura de links na rede *World Wide Web* de forma recursiva [Page et al., 1999]. O valor de PageRank de um vértice é influenciado não apenas pelo seu grau de entrada, mas também pelo valor de PageRank e do grau de saída dos vértices que apontam para ele. Com isso, a importância de um vértice depende também da importância dos vértices que apontam para ele.

A análise da estabilidade de ranqueamento em redes livres de escala foi inicialmente proposta no trabalho de Ghoshal e Barabási [Ghoshal e Barabási, 2011]. Nesse trabalho os autores afirmam que, em redes livre de escala, alguns vértices podem ser considerados superestáveis, ou seja, alguns vértices podem manter suas respectivas posições no ranqueamento por muito tempo durante o qual a rede evolui. A premissa se baseia na capacidade desses vértices manterem as suas posições no ranqueamento de PageRank. Através de simulações numéricas, os autores atestam que, em determinadas redes, as perturbações que alteram a relação de vizinhança entre os vértices sem haver alteração de grau, não alteram a ordem dos vértices no topo do ranqueamento. Isso ocorre quando a diferença entre o PageRank de um vértice superestável e do próximo colocado no ranqueamento é maior do que a variação de PageRank do vértice. A cauda pesada da distribuição de grau pode levar à ocorrência de vértices com grau muito maior do que os demais, que manterão alto PageRank mesmo que as arestas de entrada originem em vértices com baixo PageRank. Isto é, a grande diferença de grau entre os vértices de maior grau garante a estabilidade no ranqueamento desses vértices. Em contrapartida, redes com distribuição de grau exponencial não possuem vértices superestáveis por não apresentarem cauda pesada.

Além de simulação, Ghoshal e Barabási apresentam uma formulação matemática para calcular a quantidade de vértices superestáveis em função da quantidade total de vértices e do expoente da lei de potência que melhor representa sua distribuição de grau. Os resultados calculados através da equação final são aproximadamente iguais aos resultados verificados em simulações numéricas, onde redes sintéticas foram perturbadas de modo a alterar as conexões sem modificar o grau de cada vértice. As quantidades de vértices superestáveis verificadas nas redes livre de escala são pequenas em relação ao total da rede (por exemplo, são menos de dez para redes com 10^7 vértices), mas em geral aumentam com o tamanho da rede. Adicionalmente, uma análise empírica do comportamento do ranqueamento de grau de redes reais ao longo do tempo apontou que a mesma quantidade aproximada de vértices superestáveis manteve seu ranqueamento de grau por um longo tempo enquanto que os demais vértices apresentaram maior instabilidade.

Em outro trabalho, Blumm et al. [Blumm et al., 2012] analisam a dinâmica do

ranqueamento em sistemas complexos, considerando seis diferentes conjuntos de dados de sistemas reais que não representam necessariamente redes. Cada sistema é formado por um conjunto de itens e existe uma característica que é medida individualmente ao longo do tempo para ranquear os itens. Por exemplo, no sistema de citações de artigos é considerada a quantidade de citações recebidas anualmente por cada artigo (item). Repare que esse estudo considera o número de citações que cada artigo recebe em cada ano, sem acumular diferentes anos.

Em um sistema complexo, considere $a_i(t)$ a quantidade absoluta de uma característica de interesse do item i em um instante de tempo t , e $s(t)$ a soma das quantidades absolutas de todos os itens no mesmo instante. Por conseguinte, o valor relativo da característica do item i pode ser calculado como $x_i(t) = a_i(t)/s(t)$, onde $\sum_i x_i(t) = 1$. A partir da análise empírica dos seis conjuntos de dados, Blumm et al. constataram que a dispersão das variações de x a cada passo de tempo é proporcional a x^β , ou seja

$$\sigma_{\Delta x|x} = \sqrt{\frac{1}{R} \times \sum_{\forall i,t|x_i(t)=x}^R (x_i(t+1) - x_i(t))^2} \sim x^\beta, \quad (1)$$

onde R é a quantidade de ocorrências de x .

Diante da constatação dada pela Equação 1, os autores formularam uma representação matemática da dinâmica do sistema, onde o valor de x_i é influenciado não só pela aptidão individual do item em relação à qualidade medida mas também pelo ruído presente no sistema. A parcela referente ao ruído do sistema possui amplitude definida pela equação

$$g(x_i) = Bx_i^\beta, \quad (2)$$

onde $g(x_i)$ é a amplitude do ruído que afeta o item i , e os valores do coeficiente B e do expoente β são obtidos a partir de $\sigma_{\Delta x|x}$.

Por fim, os autores concluem que o valor de B está diretamente relacionado à estabilidade do ranqueamento do sistema. Ou seja, quando B é maior do que um limiar a posição dos itens no ranqueamento sofre variações causadas pelo ruído. Dentre os conjuntos de dados avaliados, os sistemas considerados estáveis apresentam valor de B na ordem de 10^{-3} enquanto que os sistemas instáveis possuem B na ordem de 10^{-2} ou 10^{-1} . Já o valor do expoente β foi menor que 1 nos sistemas avaliados, indicando que a amplitude do ruído é proporcionalmente menor nos itens mais próximos do topo do ranqueamento.

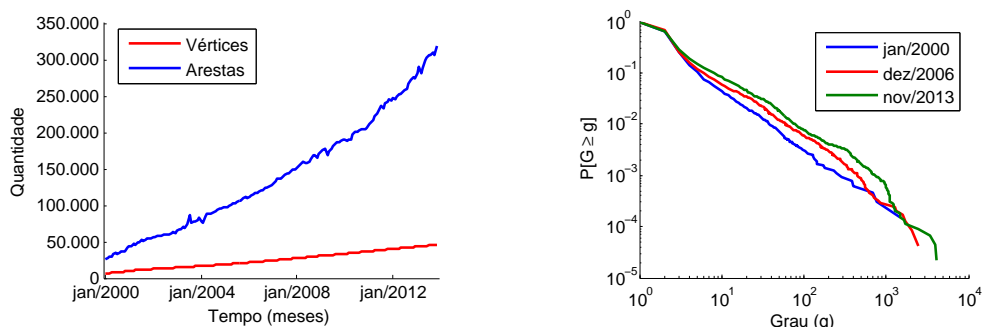
Diferente dos trabalhos anteriores, este trabalho analisa a estabilidade de ranqueamento na rede de sistemas autônomos da Internet levando em conta também o tempo de permanência em cada posição do ranqueamento e o impacto de fatores externos à estrutura da rede.

3. Conjuntos de Dados

Os conjuntos de dados contendo as listas mensais de conexões entre sistemas autônomos da Internet foram obtidos através da página *web* da CAIDA (*The Cooperative Association for Internet Data Analysis*). Os dados disponibilizados pela CAIDA foram inferidos a partir das tabelas de roteamento divulgadas na Internet através do protocolo

BGP (*Border Gateway Protocol*) [CAIDA, 2015]. Neste trabalho, as análises estão baseadas nos dados referentes ao período de janeiro de 2000 a novembro de 2013, por ser o maior período sem interrupção disponível na página da CAIDA. Os dados referentes a novembro de 2014 e novembro de 2015 também são utilizados para uma análise extra.

Neste trabalho a rede de sistemas autônomos é modelada como um grafo, onde os vértices representam os sistemas autônomos e as arestas não-direcionadas são formadas entre sistemas autônomos diretamente conectados. O grafo evolui no tempo, uma vez que o conjunto de dados possui a evolução da rede em fotografias mensais. A partir dessa mesma modelagem, diversos trabalhos observaram que a distribuição de grau da rede de sistemas autônomos se aproxima a uma lei de potência [Chen et al., 2002, Siganos et al., 2003, Ahmad e Guha, 2012]. Entre o início e o fim do período considerado neste trabalho, a rede de sistemas autônomos cresceu mais de 5 vezes em quantidade de vértices e mais de 10 vezes em número de arestas (Figura 1(a)). Apesar disso, a distribuição de grau da rede manteve sua característica livre de escala com pouca variação, conforme ilustrado na Figura 1(b) que apresenta a distribuição de grau da rede no início, meio e fim do período em análise neste trabalho.



(a) Evolução temporal da quantidade de vértices e arestas.

(b) Distribuição de grau da rede.

Figura 1. Caracterização do conjunto de dados de sistemas autônomos da Internet. Apesar do número de vértices e arestas aumentar significativamente ao longo dos anos, a distribuição de grau se mantém com pouca variação.

A Figura 2 apresenta a evolução do ranqueamento dos dez maiores graus da rede a cada mês. Cada linha com cor diferente representa um sistema autônomo diferente. Note que no período em análise, 29 sistemas autônomos diferentes ocuparam as 10 primeiras posições do ranqueamento de grau, demonstrando uma certa instabilidade geral mesmo no topo do ranqueamento. Em particular, repare que o primeiro no ranqueamento em 2000 se manteve na mesma posição por nove anos consecutivos mas ao fim do período ocupou a sexta posição, indicando que fatores pontuais influenciam a aptidão dos sistemas autônomos e conseqüentemente a habilidade em manter suas posições no ranqueamento. Por outro lado, observa-se que o tempo que os sistemas autônomos permanecem na mesma posição é maior nas primeiras posições do ranqueamento.

4. Metodologia Proposta da Análise de Estabilidade

Os fatores que influenciam a estabilidade do ranqueamento de grau na rede de sistemas autônomos são analisados neste trabalho sob três diferentes aspectos: os vértices

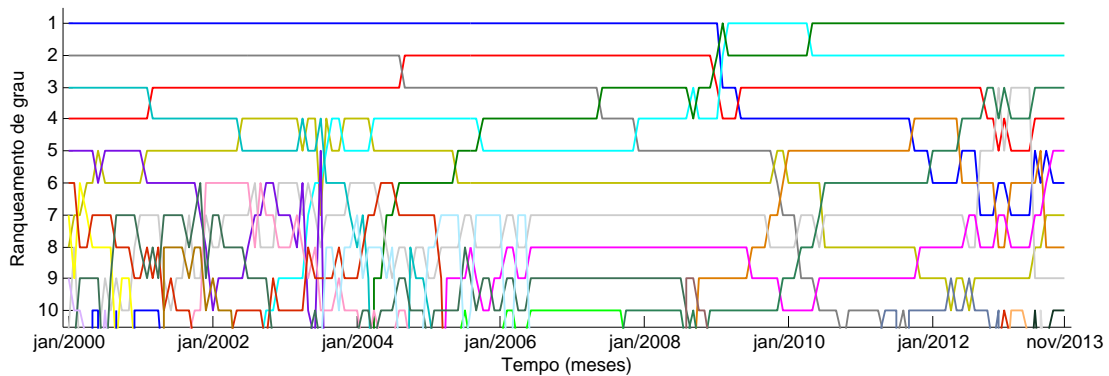


Figura 2. Variações nas dez primeiras posições do ranqueamento de grau na rede de sistemas autônomos da Internet.

superestáveis da rede, o ruído e os fatores pontuais. O objetivo final da análise é compreender como cada característica ou fator afeta o ranqueamento, de modo que seja possível também prever a dinâmica do ranqueamento.

1. **Análise da quantidade e da estabilidade temporal dos vértices superestáveis:** A quantidade prevista de vértices superestáveis (m_c) no período em análise é calculada a partir da equação proposta por Ghoshal e Barabási [Ghoshal e Barabási, 2011]. Os dados de entrada dessa equação são a quantidade total de vértices (N) e o expoente da lei de potência aproximada da distribuição de grau (γ) de uma fotografia da rede. Como essas características da rede podem sofrer alteração ao longo do tempo, neste trabalho são consideradas diferentes fotografias correspondentes a diferentes instantes de tempo. Para estimar γ a partir da distribuição de grau, foi escolhida a técnica MLE (estimador de máxima verossimilhança). Em seguida, para avaliar a relação da estabilidade com o tempo, foi introduzido neste trabalho o cálculo do tempo médio de estabilidade do ranqueamento, que é a média de tempo em que um mesmo vértice se manteve ininterruptamente em cada posição do ranqueamento. Por fim, o tempo médio de estabilidade dos vértices apontados como superestáveis é comparado com o tempo médio de estabilidade dos demais vértices.
2. **Análise do impacto do ruído no ranqueamento:** Conforme proposto por Blumm et al. [Blumm et al., 2012], o nível de estabilidade geral do ranqueamento está relacionado à amplitude do ruído, que é calculada a partir do histórico de variações de grau da rede ao longo do período em análise. Para reproduzir a técnica, mediu-se as variações de grau relativo ($\Delta x|x$) e em seguida extraiu-se o coeficiente de amplitude do ruído (B) a partir da dispersão suavizada de $\Delta x|x$, conforme proposto em [Blumm et al., 2012]. Os programas desenvolvidos foram avaliados através da análise do conjunto de dados de citações da coleção Physical Review [Society, 2015]. A estratégia de refazer os cálculos para o mesmo conjunto de dados analisado no artigo original permitiu comparar os resultados e validar o programa. Neste trabalho, além de calcular o valor de B para todo o período analisado, o mesmo cálculo é feito para períodos menores e os resultados são comparados com as variações observadas no ranqueamento de grau e nas quantidades de adições e remoções de arestas no mesmo período. O objetivo é avaliar se a

variação de B tem correlação com a variação estrutural observada empiricamente na rede.

3. **Análise do impacto de fatores pontuais no ranqueamento:** Fatores pontuais, que influenciam as variações de grau de um único vértice ou de um grupo de vértices, também afetam a estabilidade do ranqueamento da rede. São alguns exemplos de fatores pontuais investigados nesse trabalho:

- dificuldades financeiras enfrentadas pelas companhias;
- estratégias comerciais agressivas;
- estratégias comerciais restritivas;
- intervenções de órgãos reguladores para evitar concentração no mercado;
- limitações tecnológicas, como o esgotamento de endereçamento IPv4 em certas regiões.

Como as origens dos fatores pontuais são muito diversas e suas ocorrências pouco documentadas, é inviável a análise de todos os fatores pontuais de uma rede do tamanho da Internet. Neste trabalho são levantados os fatores pontuais relacionados aos principais vértices do ranqueamento de grau, isto é, aos vértices que ocuparam as primeiras posições do ranqueamento ao longo do período em análise. Após a identificação dos fatores pontuais, o impacto de cada um no ranqueamento de grau é avaliado de forma empírica, através da análise da evolução histórica do grau desses vértices.

A Figura 3 resume as etapas da metodologia empregada na análise de estabilidade do ranqueamento, destacando a contribuição deste trabalho em cada uma das etapas.

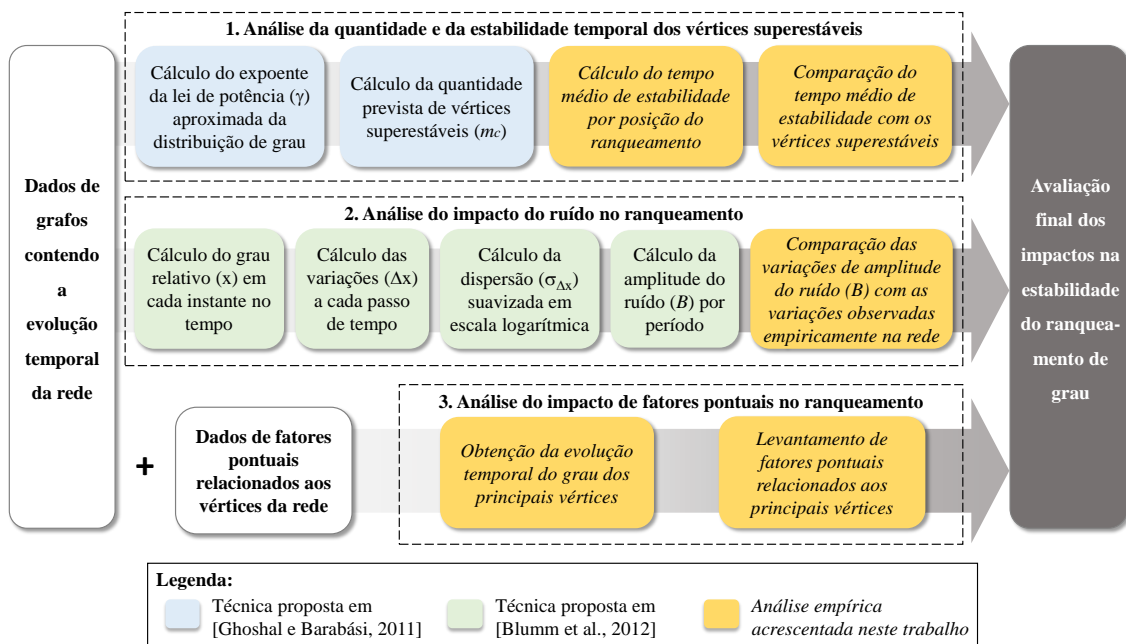


Figura 3. Metodologia de análise da estabilidade de ranqueamento. A metodologia é formada pela análise de três aspectos: os vértices superestáveis da rede, o ruído e os fatores pontuais.

5. Resultados

Nesta seção são apresentados os resultados das análises propostas na Seção 4 para a rede de sistemas autônomos da Internet. Ao final é realizada uma avaliação da previsão feita a partir dos resultados.

5.1. Análise da quantidade e da estabilidade temporal dos vértices superestáveis

A quantidade prevista de vértices superestáveis (m_c) na rede de sistemas autônomos, calculada através da equação proposta por Ghoshal e Barabási [Ghoshal e Barabási, 2011], é de aproximadamente dois vértices em todo o período analisado. Isso quer dizer que as duas primeiras posições no ranqueamento de grau da rede de sistemas autônomos devem apresentar estabilidade muito maior do que as demais posições do ranqueamento. Na Tabela 1 estão os valores de N (quantidade total de vértices) e γ (expoente da lei de potência aproximada da distribuição de grau) utilizados no cálculo de m_c , além do grau médio e dos três maiores graus, referentes ao estado da rede no oitavo mês dos anos de 2000, 2003, 2006, 2009 e 2013. É possível constatar que os três maiores graus são muito maiores do que o grau médio em todos os casos. Em 2009 os três maiores graus estavam muito próximos, sugerindo uma baixa estabilidade no ranqueamento entre eles (fato que pode ser confirmado na Figura 2). Também é observado um aumento da densidade de arestas na rede (grau médio), o que significa que a rede de sistemas autônomos está cada vez mais conectada, aumentando a disponibilidade de sistemas autônomos para trânsito na Internet.

Tabela 1. Propriedades da rede e quantidade prevista de vértices superestáveis em cinco diferentes instantes no tempo.

Rede	Tamanho N da rede	γ aproximado	m_c previsto	Grau médio	Maior grau	2º maior grau	3º maior grau
ago/2000	8.284	2,4	2	4,2	1852	996	814
ago/2003	15.821	2,3	2	4,8	2446	1808	1667
ago/2006	23.102	2,3	2	5,2	2407	2029	1740
ago/2009	32.265	2,3	2	5,7	2487	2350	2241
ago/2013	45.067	2,2	2	6,8	4042	3815	3217

A Figura 4 mostra o tempo médio de estabilidade de cada posição do ranqueamento de grau ao longo de todo o período analisado. As duas primeiras posições do ranqueamento apresentam tempo médio de estabilidade bem superior às demais posições, confirmando o resultado obtido através da equação proposta por Ghoshal e Barabási. Entretanto, ao contrário do que sugere o nome superestável, as duas primeiras posições do ranqueamento não são estáveis por todo o período analisado, pelo contrário, considerando o período total analisado de 167 meses, o tempo médio de estabilidade foi de aproximadamente 42 meses (ou 25% do tempo total) para a primeira posição e 28 meses (ou 17% do tempo total) para a segunda posição. O tempo médio de estabilidade da terceira posição foi de aproximadamente 12 meses (ou 7% do tempo total), sendo menos da metade da estabilidade observada na segunda posição. Apesar da estabilidade das duas primeiras posições do ranqueamento se destacarem, vale ressaltar também que o tempo médio apresenta uma tendência contínua de queda à medida que se observa posições mais distantes do topo do ranqueamento.

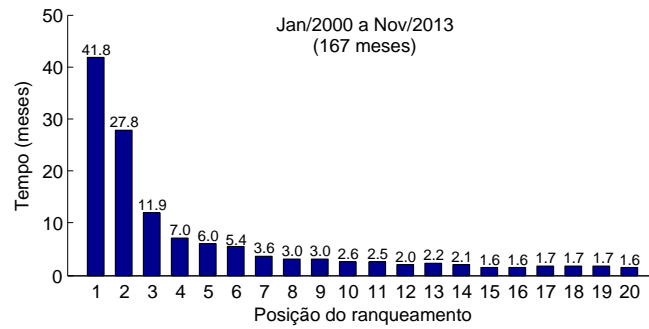
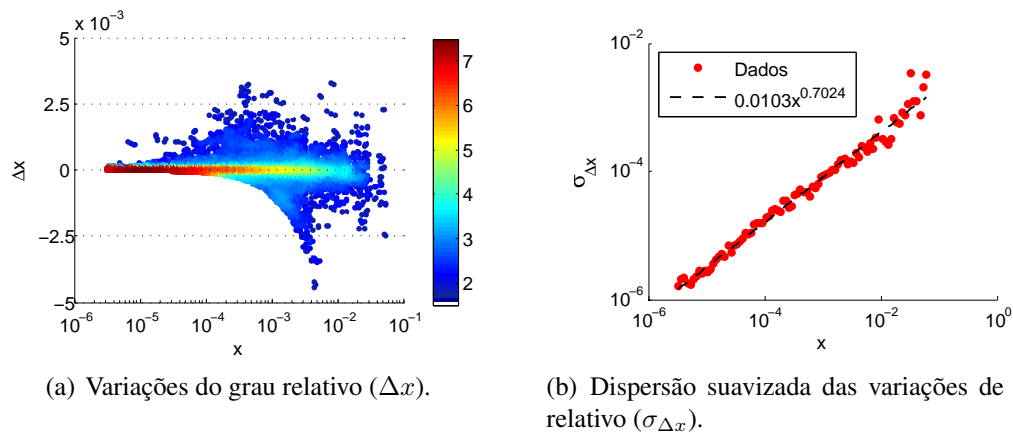


Figura 4. Estabilidade média por posição de ranqueamento na rede de sistemas autônomos.

5.2. Análise do impacto do ruído no ranqueamento

O nível de estabilidade geral do ranqueamento da rede é avaliado através da técnica desenvolvida por Blumm et al. [Blumm et al., 2012]. Para isso, foram calculadas todas as variações entre meses subsequentes dos valores relativos de grau (Δx) e em seguida a dispersão dessas variações ($\sigma_{\Delta x}$). As Figura 5(a) e 5(b) apresentam, respectivamente, os gráficos de Δx e $\sigma_{\Delta x}$ considerando todo o período em análise. Na Figura 5(a) as cores mais quentes indicam maior densidade de pontos. É possível observar uma relativa assimetria em torno de $\Delta x = 0$, onde os pontos relativos aos valores mais altos de x estão mais concentrados abaixo de $\Delta x = 0$. Isso significa que os sistemas autônomos apresentam tendência geral de variar de posição no ranqueamento de grau com o passar do tempo, sugerindo uma baixa estabilidade mesmo no topo do ranqueamento.



(a) Variações do grau relativo (Δx).

(b) Dispersão suavizada das variações de grau relativo ($\sigma_{\Delta x}$).

Figura 5. Características de variação do grau relativo na rede de sistemas autônomos.

A linha tracejada na Figura 5(b) representa a linha de tendência (i.e., regressão linear) dos valores de $\sigma_{\Delta x}$, a partir da qual é estimada a amplitude do ruído (B) presente na rede de sistemas autônomos. Nessa avaliação, considerando todo o período em análise, $B = 0,0103 \sim 1 \times 10^{-2}$. Esse resultado classifica essa rede em uma situação intermediária na classificação de estabilidade proposta por Blumm et al. Isto é, uma parcela das posições próximas ao topo do ranqueamento apresentam estabilidade enquanto que as demais posições apresentam instabilidade. Essa avaliação obtida através do valor estimado de B está coerente com a avaliação do tempo médio de estabilidade por posição do

ranqueamento mostrada na Figura 4, uma vez que as primeiras posições no ranqueamento têm um tempo médio de estabilidade bem superior do que as demais.

Uma questão interessante é a relação entre B e as variações ocorridas na rede. Tal relação foi abordada calculando o valor de B em períodos menores e comparado com as variações detectadas na rede no mesmo período. A Tabela 2 mostra, para cada período de um ano, o valor estimado da amplitude de ruído B , as quantidades proporcionais de arestas adicionadas e removidas na rede, e as quantidades de alterações nas duas primeiras posições do ranqueamento de grau. Para facilitar a visualização, foi acrescentada uma gradação de cor onde o menor valor de cada coluna é marcado de branco e o maior valor de vermelho.

Tabela 2. Amplitude do ruído e variações em cada ano.

Período	Coefficiente B do ruído	% média de arestas adicionadas por mês	% média de arestas removidas por mês	Alterações na 1ª posição do ranqueamento	Alterações na 2ª posição do ranqueamento
2000	0,017	11,4%	6,8%		
2001	0,009	9,5%	7,5%		
2002	0,006	7,2%	6,3%		
2003	0,020	8,3%	6,0%		
2004	0,008	6,5%	5,2%		1
2005	0,004	5,4%	4,3%		
2006	0,002	5,1%	3,9%		
2007	0,003	5,2%	3,9%		
2008	0,004	5,9%	4,9%		
2009	0,006	5,3%	4,4%	2	3
2010	0,003	4,6%	4,0%	1	1
2011	0,003	5,0%	3,6%		
2012	0,004	5,2%	4,3%		
2013	0,005	6,0%	4,8%		

A Tabela 2 mostra uma certa correlação entre o valor de B e a quantidade de adições e remoções de arestas. Por exemplo, no ano de 2000, o valor do ruído B é elevado, assim como a porcentagem de adições e remoções. Já em 2006, o valor de B é menor, assim como as porcentagens de adições e remoções de arestas. Por outro lado, não há uma correlação aparente do valor de B com as variações ocorridas nas duas primeiras posições do ranqueamento de grau. Por exemplo, apesar do ano de 2009 ter sido o ano com o maior número de alterações nas primeiras posições, o valor de B não foi dos mais elevados. Isso leva a crer que o ruído tem pouca ou nenhuma influência nas trocas das primeiras posições do ranqueamento dessa rede.

O ruído, da forma que foi definido por Blumm et al., causa variações aleatórias no grau relativo ($x_i(t)$) a cada passo de tempo, que no caso dessa análise é de um mês. A Figura 6 mostra a evolução do grau relativo de todos os sistemas autônomos que ocuparam a primeira e a segunda posições do ranqueamento de grau no período em análise. É possível identificar um ruído de maior amplitude ao longo dos anos de 2000 e de 2003, coerente com os picos do valor estimado de B mostrados na Tabela 2. Note que entre 2000 e 2008 a diferença de grau relativo dos dois primeiros colocados é grande, maior do que os picos de variação causados pelo ruído. A partir de 2009, a diferença entre os dois primeiros colocados se mantém pequena e passam a ocorrer algumas trocas na primeira

posição. Essa condição de proximidade entre o grau da primeira e da segunda posições é inesperada em uma distribuição de grau livre de escala, indicando que a evolução de grau desses vértices sofre influência de fatores pontuais.

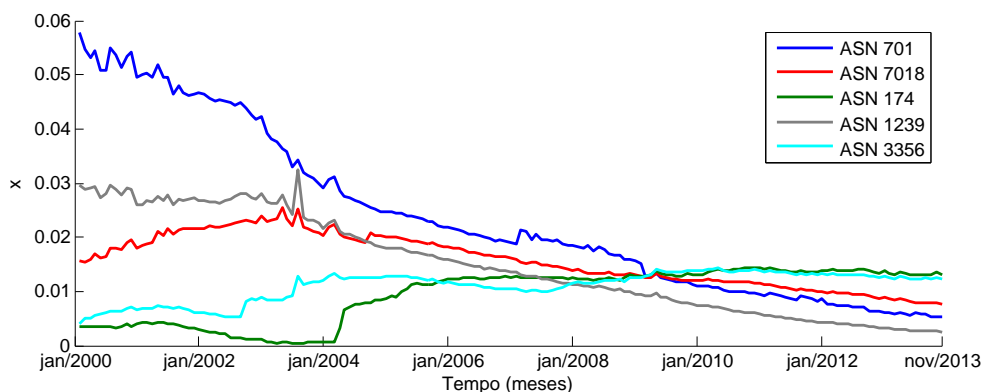


Figura 6. Evolução do grau relativo (x_i) dos principais sistemas autônomos.

5.3. Análise do impacto de fatores pontuais no ranqueamento

Conforme a premissa adotada por Blumm et al. [Blumm et al., 2012], o valor de x_i é determinado não apenas pelo ruído presente no sistema, mas também pela aptidão individual de i . Analogamente, a estabilidade de ranqueamento de cada sistema autônomo depende também da sua capacidade de atrair conexões. Se um sistema autônomo aumentar sua capacidade de atrair conexões em relação aos demais, ele tende a subir de posição no ranqueamento. Se reduzir essa capacidade, tende a perder posição. Apesar de prever a influência da aptidão individual de i no cálculo de x_i , Blumm et al. não definiram uma forma de calcular a aptidão.

Os resultados até aqui indicam que, devido às características da rede, as duas primeiras posições do ranqueamento de grau possuem elevada estabilidade e o nível de ruído do sistema não é determinante para causar as variações observadas nessas posições. Resta, portanto, avaliar se há fatores pontuais responsáveis por alterar a aptidão dos principais sistemas autônomos de modo a influenciar as variações no topo do ranqueamento.

No período em análise, as duas primeiras posições do ranqueamento de grau foram ocupadas pelos sistemas autônomos de ASN (*Autonomous System Number*) 174, 701, 1239, 3356 e 7018. Esses ASNs pertencem a companhias privadas, portanto a expansão das redes desses sistemas autônomos depende fundamentalmente da estratégia de crescimento e da capacidade de investimento de cada companhia. A seguir são listados os fatos relevantes a respeito dessas companhias divulgados publicamente dentro do período analisado:

- A) Em 2002 a companhia detentora do ASN 701 (MCI/WorldCom) entrou em concordata [CNN, 2002]. A companhia continuou passando por dificuldades nos anos seguintes, culminando na venda da companhia em 2006 para uma concorrente (Verizon) [Verizon, 2007]. Neste período o grau do ASN 701 encerrou a tendência anterior de crescimento e começou a recuar. Houve uma redução da aptidão desse sistema autônomo.

- B) Em 2004 ocorreu a expansão da empresa detentora do ASN 174 (Cogent) para a Europa e a taxa de crescimento do seu grau aumentou significativamente [Cogent, 2015]. Nesse caso ocorreu um aumento da aptidão do sistema autônomo.
- C) Entre 2004 e 2005 houve a fusão da companhia detentora do ASN 1239 (Sprint) com uma companhia operadora de telefonia móvel (Nextel). Essa fusão foi considerada problemática, dentre outros motivos, por causa da disparidade das tecnologias das duas redes [Forbes, 2012]. A concentração de esforços na integração das redes de telefonia móvel nos anos seguintes à fusão causou redução de investimentos na rede do sistema autônomo, levando à redução da sua aptidão.
- D) Entre 2004 e 2007 a companhia detentora do ASN 3356 (Level 3) adquiriu diversas empresas concorrentes [Wikipedia, 2015]. A partir de 2007 a taxa de crescimento do grau aumentou significativamente como resultado da estratégia de expansão, levando ao aumento da aptidão desse sistema autônomo.

Os órgãos reguladores governamentais têm poder de limitar a atuação das maiores companhias do mercado para estimular a competição, como ocorreu com o sistema Bell na década de 1940. Contudo não foi identificado impacto desse tipo no período em análise. Outro fato relevante é o esgotamento recente do endereçamento IPv4 mas não há impactos evidentes desse tipo nas redes dos cinco principais sistemas autônomos no período em análise.

A Figura 7 mostra a evolução do grau dos principais sistemas autônomos com a indicação da atuação dos fatores pontuais. Nessa figura estão os mesmos sistemas autônomos mostrados na Figura 6, porém com os valores absolutos de grau ao invés dos valores relativos. É possível observar que o fator determinante para a definição das duas primeiras posições do ranqueamento de grau é a taxa de crescimento de grau de cada sistema autônomo, que está associada à aptidão individual do sistema autônomo. Conforme analisado na seção anterior, o ruído apresentou amplitude insuficiente para causar troca nas duas primeiras posições do ranqueamento, com exceção de um pequeno período entre 2009 e 2010 onde o grau dos principais sistemas autônomos esteve muito próximo e o ruído pode ter influenciado momentaneamente o ranqueamento entre eles.

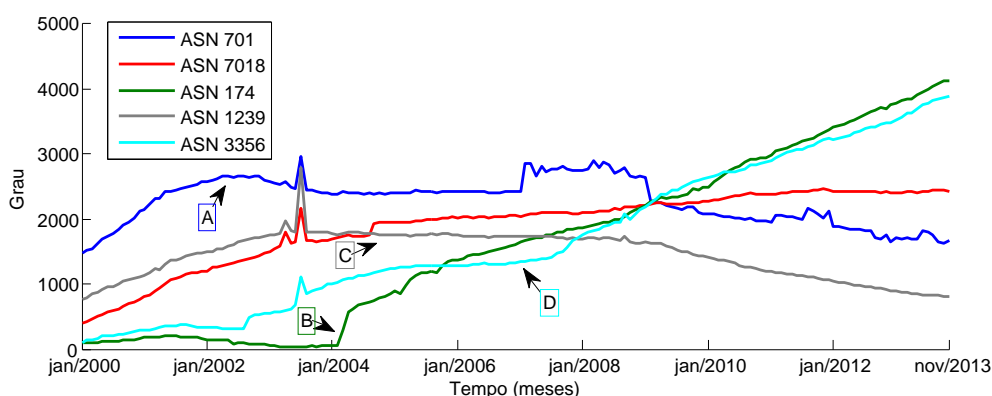


Figura 7. Evolução do grau dos principais sistemas autônomos. As marcações destacam as ocorrências dos fatores pontuais.

5.4. Avaliação da previsão de estabilidade do ranqueamento

Diante dos resultados obtidos, é esperado que as duas primeiras posições do ranqueamento de grau da rede de sistemas autônomos continuem com alta estabilidade em

relação às demais posições. Essa estabilidade será mantida até que ocorram alterações nas aptidões dos principais vértices da rede de forma a causar uma troca nessas posições. E de fato os mesmos sistemas autônomos continuam a ocupar as duas primeiras posições do ranqueamento de grau após dois anos desde o fim do período analisado neste trabalho. A Tabela 3 mostra o ASN e o grau dos cinco sistemas autônomos de maior grau no último mês do período analisado e nos dois anos seguintes, com os dois ASN de maior grau destacados em vermelho. A terceira posição do ranqueamento também foi ocupada pelo mesmo sistema autônomo nos dois anos seguintes, porém a quarta e quinta posições apresentaram variações conforme esperado. Nota-se que o crescimento do grau dos três maiores sistemas autônomos se manteve regular nesse período, indicando que não houve mudança na aptidão desses sistemas autônomos, garantindo a estabilidade do ranqueamento.

Tabela 3. ASN e grau dos 5 sistemas autônomos de maior grau ano final do período analisado e após 1 e 2 anos.

Ranqueamento de grau	Nov/2013		Nov/2014		Nov/2015	
	ASN	Grau	ASN	Grau	ASN	Grau
1º	174	4137	174	4306	174	4765
2º	3356	3897	3356	3990	3356	4284
3º	6939	3408	6939	3598	6939	4270
4º	7018	2433	3549	3573	3549	3560
5º	4323	1710	7018	2340	24482	2694

6. Conclusões e Trabalhos Futuros

A cauda pesada na distribuição de grau da rede de sistemas autônomos é uma característica que se mantém constante ao longo do tempo. Com isso, o ranqueamento por grau dos vértices leva a grandes diferenças de grau no topo do ranqueamento. Consequentemente, as trocas de posição próximas ao topo do ranqueamento são mais raras do que nas posições mais distantes do topo.

As variações regulares do grau relativo, que podem ser comparadas a um ruído, causam relativa instabilidade no ranqueamento de grau da rede de sistemas autônomos. Porém, como analisado neste trabalho, a estabilidade dos sistemas autônomos nas duas primeiras posições do ranqueamento não é afetada pelo ruído, dado que a sua amplitude é muito menor do que a amplitude da distância média entre os três vértices de maior grau. As variações observadas nas duas primeiras posições foram causadas por fatores pontuais, que alteraram a capacidade individual dos principais sistemas autônomos de atrair conexões, ou em outras palavras, a aptidão. Sendo assim, os fatores aqui chamados de pontuais devem ser levados em conta na análise de estabilidade, principalmente de sistemas autônomos, algo que não foi considerado em trabalhos anteriores.

Como trabalho futuro, conjuntos de dados de outras redes livre de escala serão analisados para confirmar a aplicabilidade global da metodologia utilizada neste trabalho. Em particular, da influência de fatores pontuais na mudança de aptidão dos vértices e consequente instabilidade no ranqueamento.

Referências

Ahmad, M. Z. e Guha, R. (2012). Studying the effects of Internet exchange points on Internet topology. *J Inform Tech Softw Eng*, 2:114.

- Albert, R., Jeong, H. e Barabási, A.-L. (1999). Diameter of the world wide web. *Nature Communications*, 2(6749):130–131.
- Barabási, A.-L. e Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.
- Blumm, N., Ghoshal, G., Forró, Z., Schich, M., Bianconi, G., Bouchaud, J.-P. e Barabási, A.-L. (2012). Dynamics of ranking processes in complex systems. *Physical Review Letters*, 109(12):128701.
- Boldi, P. e Vigna, S. (2014). Axioms for centrality. *Internet Mathematics*, 10(3-4):222–262.
- CAIDA (2015). The CAIDA AS relationships dataset. Disponível em <http://data.caida.org/datasets/2013-asrank-data-supplement/>.
- Chen, Q., Chang, H., Govindan, R. e Jamin, S. (2002). The origin of power laws in Internet topologies revisited. Em *INFOCOM*, p. 608–617 vol.2.
- CNN (2002). WorldCom files largest bankruptcy ever. Acessado 11/2015 em http://money.cnn.com/2002/07/19/news/worldcom_bankruptcy/.
- Cogent (2015). Cogent history. Acessado 11/2015 em <http://www.cogentco.com/en/about-cogent/history>.
- Figueiredo, D. R. (2011). Introdução a redes complexas. Em *Atualizações em Informática*, chapter 7, p. 303–358. PUC-Rio.
- Forbes (2012). Was Sprint buying Nextel one of the worst acquisitions ever at \$35b? Acessado 11/2015 em <http://www.forbes.com/sites/quora/2012/11/29/was-sprint-buying-nextel-one-of-the-worst-acquisitions-ever-at-35b/>.
- Ghoshal, G. e Barabási, A.-L. (2011). Ranking stability and super-stable nodes in complex networks. *Nature Communications*, 2(394):1–7.
- Page, L., Brin, S., Motwani, R. e Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Relatório Técnico 1999-66, Stanford InfoLab.
- Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, 4(2):131–134.
- Siganos, G., Faloutsos, M., Faloutsos, P. e Faloutsos, C. (2003). Power laws and the AS-level Internet topology. *IEEE/ACM Transactions on Networking*, 11(4):514–524.
- Society, A. P. (2015). APS data sets for research. Disponível em <http://journals.aps.org/datasets>.
- Verizon (2007). Verizon Communications 2007 annual report. Acessado 11/2015 em https://www.verizon.com/investor/app_resources/interactiveannual/2007/note08.html.
- Wikipedia (2015). Level 3 — Wikipedia, the free encyclopedia. Acessado 11/2015 em https://en.wikipedia.org/wiki/Level_3_Communications.