

Anatomia do Ecossistema de Pontos de Troca de Tráfego Públicos na Internet do Brasil

Samuel Henrique Bucke Brito¹, Mateus Augusto Silva Santos¹,
Ramon dos Reis Fontes¹, Danny Alex Lachos Perez¹,
Christian Esteve Rothenberg¹

¹ Universidade Estadual de Campinas (UNICAMP)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Information & Networking Technologies Research & Innovation Group (INTRIG)
Av Albert Einstein, 400, Cidade Universitária Zeferino Vaz, Campinas, SP, Brasil

{shbbrito, msantos, ramonrf, dlachosp, chesteve}@dca.fee.unicamp.br

Abstract. *For a long time many efforts were carried to increase the understanding of the complex Internet ecosystem. Internet eXchange Points have been targets since they represent a microcosm of Internet diversity. The contribution of this paper is the first empirical analysis of the IXP ecosystem in Brazil, compiling information that comprehend the characterization of members participating of these environments, including the build of AS-level connectivity graphs. By analysing the topologies we found the average adjacency of IXP's members (vertices degree), the isolated and accumulated depth of paths in advertised routes (AS-PATH), the density of peers to highlight the peering potential within IXPs, the accounting of traffic engineering policies related to AS-Prepend and k-clique communities. Another contribution is that our dataset with more than 2.5 GB is shared with the academic community.*

Resumo. *Há vários anos esforços foram conduzidos na busca por uma melhor compreensão do complexo ecossistema da Internet. Os Pontos de Troca de Tráfego (PTT) têm sido alvos constantes de pesquisas dessa natureza porque representam um microcosmo da diversidade da Internet. Este trabalho traz como contribuição a primeira análise do ecossistema de PTTs em operação no Brasil, compilando informações que compreendem desde a caracterização dos tipos de membros desses ambientes até a construção dos respectivos grafos de conectividade (em nível de Sistema Autônomo). A partir das topologias foi possível identificar a quantidade média de adjacências dos membros dos PTTs (grau dos vértices), a profundidade isolada e acumulada dos caminhos que compõem as rotas anunciadas (AS-PATH), a densidade de pares que determina o potencial de peering nos PTTs, a ocorrência de políticas de engenharia de tráfego do tipo AS-Prepend e comunidades k-clique. Outra contribuição é que o dataset com mais de 2,5 GB de informações levantadas para realização desta análise é compartilhado com a comunidade acadêmica.*

1. Introdução

Uma forma de promover o desenvolvimento da Internet em termos de conectividade e desempenho é através da implantação dos chamados Pontos de Troca de Tráfego (PTT),

do inglês *Internet eXchange Point* (IXP). Por possuírem dezenas ou centenas de Sistemas Autônomos (SA) conectados, atualmente os PTTs têm um papel crítico no ecossistema da Internet. O OpenIX¹ e o Euro-IX² são exemplos de iniciativas dos Estados Unidos e da Europa, respectivamente, que promovem o desenvolvimento de PTTs. Outro exemplo de destaque é o AMS-IX³, localizado na Holanda, que possui mais de 600 membros e atualmente é o maior PTT do mundo. No cenário nacional o projeto PTTMetro⁴ (ou PTT.br) contempla todos os PTTs públicos em operação, sendo o PTT de São Paulo (PTT-SP) o maior do Brasil (e da América Latina) com registros médios de troca de tráfego da ordem de 300 Gbps e picos de 500 Gbps. Atualmente o PTT.br está entre os dez que mais trocam tráfego no mundo, sendo o quinto maior em número de participantes.

Muitas empresas justificam seus interesses em *peering* nos PTTs pelos benefícios de desempenho e segurança [Norton 2014]. Em termos de desempenho, sabe-se que o tráfego de vídeo já representa 50% de todo o tráfego da Internet e continua crescendo, o que faz o *peering* em PTTs ideal para melhorar a distribuição desse conteúdo nas proximidades dos clientes finais, já que melhora a latência ao minimizar o diâmetro entre os SAs. Durante períodos de congestionamento no tráfego de trânsito decorrente de ataques de negação de serviço, o tráfego direto através de *peering* em PTTs é separado e não sofre das vulnerabilidades do tráfego de trânsito, o que implica em mais segurança.

Exceto para aqueles que participam do ecossistema dos PTTs, é pouco disseminado o conceito de *peering* remoto, prática que está se tornando cada vez mais comum porque permite que grandes empresas e provedores de conteúdo possam ser membros dos PTTs regionais sem os altos custos dos roteadores, do espaço físico dos equipamentos (*colocation space*) e de recursos humanos. Esse crescimento implica em mais membros nos PTTs que, por sua vez, implica em mais troca de tráfego e mais rotas no ecossistema, uma dinâmica que vem aumentando o valor estratégico dos PTTs.

Nos últimos anos foram realizados vários esforços na busca por uma melhor compreensão do complexo ecossistema da Internet [Haddadi and Bonaventure 2013], principalmente através de análises baseadas em dados públicos como: (i) informações obtidas em entrevistas ou bancos de dados, (ii) ferramentas do plano de dados (ex.: *traceroute*) e (iii) tabelas BGP do plano de controle disponíveis em servidores de rotas [Giotsas et al. 2014, Lodhi et al. 2014, Schlinker et al. 2014, Giotsas et al. 2013, Luckie et al. 2013, Ager et al. 2012, Labovitz et al. 2010]. Os PTTs têm sido alvos de pesquisas dessa natureza porque representam um microcosmo da diversidade da Internet [Chatzis et al. 2013], tendo membros que contemplam as grandes operadoras de trânsito, provedores de acesso, provedores de conteúdo, instituições públicas e empresas privadas.

No entanto esses trabalhos focaram em um único PTT de grande porte ou em poucos. Esse artigo é inovador por dois motivos principais: (i) traz um estudo do ecossistema de PTTs focado exclusivamente na realidade do Brasil e (ii) trata-se da primeira análise envolvendo todos os PTTs que existem em operação em um país de tamanho continental. Esse estudo focado no Brasil é conveniente de ser feito porque no país existe uma política pública nacional de gestão integrada dos PTTs, sob responsabilidade do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), a autoridade nacional vinculada à autoridade regional da Internet na América Latina e Caribe (LACNIC).

¹ <http://www.open-ix.org> ² <https://www.euro-ix.net> ³ <https://ams-ix.net> ⁴ <http://ptt.br>

Esse trabalho traz como principal contribuição a primeira análise do ecossistema de PTTs em operação no Brasil, compilando informações que compreendem desde a caracterização dos tipos de membros desses ambientes até a construção dos respectivos grafos de conectividade (em nível de SA) dos PTTs públicos do Brasil. A partir das topologias foi possível identificar a quantidade média de adjacências dos membros dos PTTs (grau dos vértices), a profundidade isolada e acumulada dos caminhos que compõem as rotas anunciadas (AS-PATH), a densidade de pares que determina o potencial de *peering* nos PTTs, a ocorrência de políticas de engenharia de tráfego (AS-Prepend) e a identificação de comunidades do tipo k-clique.

Outra contribuição importante é que o *dataset* com mais de 2,5 GB de informações levantadas para realização dessa análise é compartilhado com a comunidade acadêmica interessada em estudos dessa natureza e pode ser acessado através da página do nosso grupo de pesquisa⁵. Nesse amplo *dataset* a comunidade acadêmica pode encontrar a relação dos 26 PTTs do projeto PTTMetro e caracterização dos participantes, além de planilhas, grafos, gráficos e scripts elaborados no processo de análise dos dados.

2. Background

Desde 1987 até 1994 a operação do backbone da rede NSFNET era responsabilidade da *National Science Foundation* (NSF) que, a partir de 1992, deu início a um plano para transferir a operação do núcleo da Internet para o setor privado. Foi no contexto desse novo ecossistema comercial que surgiram três elementos importantes: (i) *Network Service Providers* (NSP), responsáveis pela operação do backbone; (ii) *Network Access Points* (NAP), para transportar tráfego entre os NSPs a partir de locais espalhados nos EUA; e (iii) *Routing Arbiter* (RA), para coletar e propagar informações de roteamento nos NAPs (similar aos modernos *Route Servers*). A concepção dos PTTs nasceu ainda nessa época, uma vez que os NAPs foram criados para ser um ponto físico de conexão de vários NSPs. Ao longo dos anos os NAPs enfraqueceram porque eram mantidos por grandes operadoras com interesses próprios, o que motivou a desconexão das demais operadoras.

As grandes operadoras somente tinham interesse de se conectar diretamente (*peering*) com outras de mesmo nível, demonstrando pouco (ou nenhum) interesse no *peering* aberto com outras operadoras menores. Em 1999 a necessidade por circuitos ponto-a-ponto entre as grandes operadoras escalava linearmente e custava caro, além do fato de que algumas vezes as concessionárias telefônicas levavam mais de um ano para entregar um circuito [Norton 2014]. Diante dessa situação as grandes operadoras perceberam que o estabelecimento de *peering* privado através de um PTT neutro era mais rápido e econômico, tornando esse modelo dominante na Internet atual.

Por definição o PTT é uma infraestrutura compartilhada que é instalada em uma região para receber a conexão de SAs - através de *peering* - com o objetivo principal de otimizar o desempenho da Internet ao manter a troca de tráfego mais localizada entre diferentes redes pertencentes a uma mesma região, diminuindo, assim, o número de saltos entre SAs próximos uns dos outros. Seu núcleo é bastante complexo em termos de quantidade de conexões porque hospeda centenas de membros, por isso requer equipamentos de alto desempenho capazes de processar altas taxas de pacotes por segundo (pps). Apesar dessa complexidade, sua arquitetura é simples de entender, uma vez que o

⁵ <https://github.com/intrig-unicamp/ixp-ptt-br>

objetivo do PTT se resume em prover um ponto centralizado de conexão através de uma *switching-fabric* baseada em tecnologia Ethernet (camada 2).

Uma vez fisicamente conectados no PTT, seus membros podem acordar em fazer o *peering* multilateral (aberto) com todos os demais membros ou o *peering* bilateral (privado) de natureza seletiva ou restritiva, sendo que as configurações para anúncio e alcançabilidade de prefixos IP são realizadas por meio do protocolo *Border Gateway Protocol* (BGP) [RFC4271], via sessão TCP na porta 179. Para minimizar a complexidade de configuração individual do *peering* entre todos os membros através de uma topologia *full-mesh*, são instalados elementos centrais denominados servidores de rotas (RS) na infraestrutura do PTT, de maneira que um SA é capaz de trocar rotas com os demais membros através do estabelecimento de uma única sessão BGP com os RS (*peering* multilateral), conforme pode ser observado na figura 1(a).

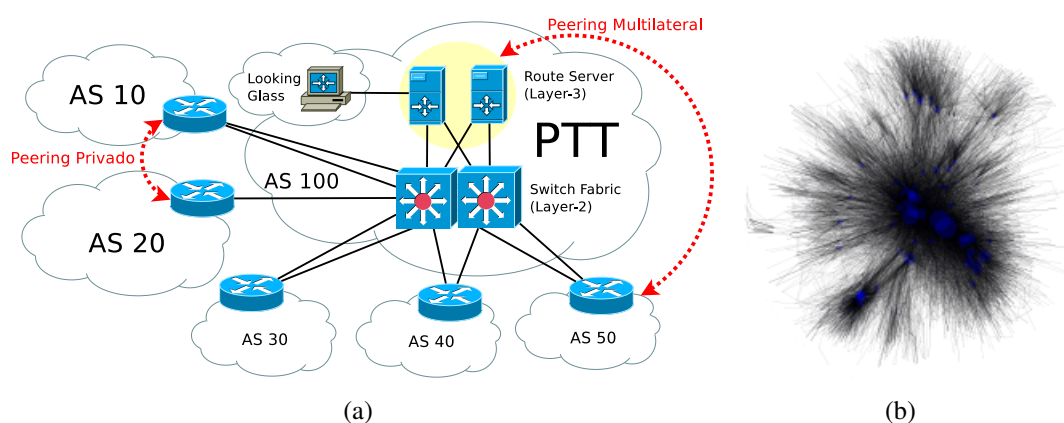


Figura 1. (a) Arquitetura do PTT (b) Exemplo de Grafo do PTT-VIX (Vitória/ES)

O modelo de negócios do PTT pode ser de natureza privada - comum nos Estados Unidos - ou aberta, como acontece na Europa e no Brasil. Na maioria dos casos, inclusive no Brasil, o objetivo principal do PTT é prover um ambiente aberto e indiscriminatório de natureza pública para estimular a colaboração e melhorar a troca de tráfego, alavancando a qualidade da Internet na sua região de operação.

2.1. Trabalhos Relacionados

Uma análise temporal interessante é apresentada em [Ager et al. 2012], baseada em nove meses de registros coletados de um dos maiores PTTs do mundo, sendo que os autores classificaram a diversidade de membros conectados. A principal descoberta desse estudo foi que a quantidade de *peerings* de um único PTT em 2012 excedeu a quantidade total de *peering* entre SAs de toda a Internet no ano de 2010.

Um estudo focado no *Slovak Internet eXchange* (SIX) é encontrado em [Restrepo and Stanojevic 2012], sob o argumento de que existem muitos esforços e medições com interesse em compreender a Internet como um todo, mas que pouco é conhecido sobre a Internet no contexto local de uma determinada região. Os autores fizeram a caracterização da evolução do ecossistema de provedores conectados no PTT da Eslováquia e do perfil de tráfego distribuído entre seus membros.

Em [Giotsas et al. 2013] é descrito um método para descoberta de *peerings* ocultos em PTTs através da mineração do atributo *community* do BGP, problema que também

é abordado em [Chen and et al 2009]. O método é particularmente interessante porque um AS pode restringir rotas anunciadas, fazendo com que a relação de AS-PATH (caminho BGP) dos *route servers* de um PTT sejam incompletas. Através da técnica os autores inferiram 206.000 *links* ponto-a-ponto a partir de 13 PTTs da Europa, número que representa quatro vezes mais *links* do que pode ser diretamente observado nas tabelas BGP.

[Giotsas et al. 2014] argumentam que tradicionalmente as relações entre SAs são classificadas em: (i) trânsito (*provider-to-customer*), (ii) *peering* (*peer-to-peer*) e (iii) *sibling* (domínios “irmãos”), no entanto existem configurações avançadas que pressupõem relações híbridas ou complexas. Para compreender essas relações complexas foi expandido o algoritmo de inferência do projeto CAIDA⁶ para que sua lógica pudesse analisar tabelas BGP, saídas do aplicativo *traceroute* e dados de geolocalização. Através desse novo algoritmo foi observado que 4,5% de um universo de 90.272 relações do tipo trânsito (*provide-to-customer*) eram, na realidade, relações híbridas ou complexas.

Em [Lodhi et al. 2014] foi realizado um trabalho recente de mineração de dados no *PeeringDB*⁷, uma das poucas fontes públicas carregadas com dados do ecossistema de *peering* e que consiste em uma ferramenta *online* em que as redes participantes da Internet contribuem inserindo informações sobre suas políticas, volumes de tráfego e presença geográfica. Nesse estudo os autores encontraram correlações consistentes entre os dados do *PeeringDB* com medidas realizadas a partir de prefixos BGP anunciados na Internet.

Em [Gregori et al. 2011] é aplicado um estudo em grafos de SAs conectados em PTTs em que são definidas como comunidades principais aquelas que possuem maiores comunidades, sendo que os autores concluem que tais comunidades são tipicamente grandes em tamanho e possuem baixa densidade de conexão, tendendo a se conectar em nós externos à comunidade. Um estudo similar utilizando comunidades *k-dense* é encontrado em [Orsini et al. 2013], cuja conclusão mostra que comunidades com o valor máximo de *k* tendem a ser compostas por provedores *tier-2* ou provedores de conteúdo.

3. Metodologia

3.1. Fontes de Informação

Duas das fontes iniciais de informações advêm dos registros do projeto PTTMetro do NIC.br e da ferramenta *online* PeeringDB. Durante o processo de análise dos dados obtidos descobrimos que o PeeringDB contém informações desatualizadas acerca dos PTTs brasileiros e seus membros, o que ficou evidente ao confrontarmos seus dados com aqueles oficiais fornecidos pelo PTT.br.

A fonte mais relevante foi a compilação de um amplo *dataset* com mais de 2,5 GB de dados a partir de um trabalho exaustivo de acesso *telnet* a todos os *Looking Glasses* (LG)⁸ dos PTTs brasileiros para coletar os seguintes dados: (i) a tabela de rotas BGP do plano de controle, (ii) relação de AS-PATH do BGP e (iii) códigos *communities*. Em 12 dos 26 PTTs **não** foi possível extrair a tabela de rotas BGP, fato confirmado pelo NIC.br, pois o acesso público à tabela BGP completa foi desativado por problemas de escalabilidade, já que os LGs não suportam a carga de consultas. Para contornar esse problema, solicitamos esses dados diretamente ao NIC.br e fomos atendidos.

⁶ www.caida.org ⁷ <http://www.peeringdb.org>

A tabela 1 traz a relação dos PTTs no Brasil, destacando suas regiões de operação, a média de troca de tráfego em Gbps, a quantidade de membros participantes e os endereços para acesso público aos seus LGs, além de observações que destacam alguns problemas que enfrentamos na coleta dos dados.

Tabela 1. PTTs Públicos em Operação no Brasil (PTTMetro)

-	Cidade	Estado	Código	Looking Glass	Gbps	Membros	Obs
01	Americana	SP	AME	lg.ame.ptt.br	0,30	9	
02	Belém	PA	BEL	lg.bel.ptt.br	1,66	13	
03	Belo Horizonte	MG	MG	lg.mg.ptt.br	3,02	34	
04	Brasília	DF	DF	lg.df.ptt.br	2,98	24	
05	Campina Grande	PB	CPV	lg.cpv.ptt.br	0,20	11	
06	Campinas	SP	CAS	lg.cas.ptt.br	2,58	33	(*)
07	Cuiabá	MT	CGB	lg.cgb.ptt.br	0,00	7	(*)
08	Caxias do Sul	RS	CXJ	lg.cxj.ptt.br	0,07	5	(*)
09	Curitiba	PR	PR	lg.pr.ptt.br	14,59	58	(1)
10	Florianópolis	SC	SC	lg.sc.ptt.br	1,05	32	(*)
11	Fortaleza	CE	CE	lg.ce.ptt.br	0,87	26	
12	Goiania	GO	GYN	lg.gyn.ptt.br	0,38	21	
13	Lajeado	RS	LAJ	lg.laj.ptt.br	0,00	7	(*)
14	Londrina	PR	LDA	lg.lda.ptt.br	1,45	31	
15	Manaus	AM	MAO	lg.mao.ptt.br	0,00	7	(*)
16	Maringá	PR	MGF	lg.mgf.ptt.br	0,06	14	(*)
17	Natal	RN	NAT	lg.nat.ptt.br	0,74	12	(*)
18	Porto Alegre	RS	RS	lg.rs.ptt.br	6,30	100	
19	Recife	PE	PE	lg.pe.ptt.br	0,03	13	
20	Rio de Janeiro	RJ	RJ	lg.rj.ptt.br	19,83	52	(*)
21	Salvador	BA	BA	lg.ba.ptt.br	1,78	41	(*)
22	São Carlos	SP	SCA	lg.sca.ptt.br	0,00	3	(*)
23	São José dos Campos	SP	SJC	lg.sjc.ptt.br	0,47	11	
24	São José do Rio Preto	SP	SJP	lg.sjp.ptt.br	0,03	8	(*)
25	São Paulo	SP	SP	lg.sp.ptt.br	334,27	550	(1)
26	Vitória	ES	VIX	lg.vix.ptt.br	0,22	20	

(1) Há filtros no LG que estão limitando a quantidade de rotas na tabela BGP.

(*) PTT com acesso indisponível ao LG, cujos dados foram fornecidos pelo NIC.br.

3.2. Construção dos Grafos

Através das ferramentas NetworkX⁹ e Neo4j¹⁰ (banco de dados orientado a grafos), os grafos de conectividade dos PTTs foram reproduzidos tendo como entradas as matrizes de adjacência geradas a partir dos arquivos extraídos dos PTTs, e.g., figura 1(b). Os grafos devem ser interpretados tendo como referência de vértices os SAs dos caminhos BGP (AS-PATH), cujas arestas representam sessões BGP inter-domínios.

Para reprodução dos grafos dos PTTs personalizamos vários algoritmos para a ferramenta NetworkX (baseados na linguagem Python) e escrevemos um algoritmo em Java para fazer a inserção dos AS-PATH BGP no Neo4j. Em relação ao processamento dos grafos no NetworkX e no Neo4j, cabe destacar que em ambas as ferramentas o processamento dos dados consumiu em média 2 horas para cada PTT, sendo que foram utilizados

⁸ Looking Glass (LG) é uma máquina pública que espelha todas as rotas existentes na tabela BGP dos membros dos PTTs. São comumente acessados via *telnet* pelos membros do PTT para fins de verificação de erros e validação de novas configurações. ⁹ <https://networkx.github.io/> ¹⁰ <http://neo4j.com/>

computadores com processador Intel Core I7-4790 3.6GHz x8 e com 16GB RAM. Por outro lado, depois de processados os dados de entrada, o NetworkX e o Neo4j permitiram executar as consultas de forma simples e rápida.

Através dos grafos foi possível aplicar análises mais aprofundadas sobre a dinâmica do ecossistema de PTTs na Internet do Brasil, cujos resultados são apresentados na próxima seção com a descrição das métricas. Os grafos individuais de todos os PTTs, suas ferramentas complementares (algoritmos) e as análises detalhadas fazem parte dos arquivos do *dataset* que estamos disponibilizando publicamente.

4. Análise e Discussão dos Resultados

Nesta seção discutimos os resultados obtidos a partir das análises dos dados coletados, compreendendo os seguintes estudos: (i) classificação dos membros dos PTTs; (ii) grau dos vértices; (iii) profundidade isolada e acumulada; (iv) densidade de pares; (v) ocorrência de AS-Prepend; e (vi) comunidades k-clique. Por limitação de espaço apresentamos apenas os resultados de quatro PTTs sem filtros: DF, MG, RS e VIX.

Nos estudos tabelados optamos por inserir uma coluna denominada Brasil equivalente à média dos grafos de todos os PTTs com seus devidos valores de desvio padrão. O alto valor observado nos desvios padrões mostra que há diversidade no panorama nacional, sendo importante destacar que a diferença entre PTTs maiores e menores também influencia nessa dispersão. Por exemplo, em PTTs menores que possuem menos membros é normal que haja maior grau de conexão (densidade) porque há menos combinações possíveis, embora essa característica não possa ser assumida como regra.

4.1. Classificação dos Membros

Um primeiro esforço de organização dos trabalhos de análise foi classificar todos os SAs dos PTTs brasileiros, tarefa que nos permitiu identificar o perfil atual dos membros interessados em *peering* para troca de tráfego nas diversas regiões do Brasil. No *dataset* compartilhado pode ser encontrada a classificação individual de todos os PTTs, enquanto que na tabela 2 trazemos uma síntese com o resultado da classificação geral do cenário nacional e algumas classificações individuais, incluindo o PTT-SP.

Ao observar os resultados da tabela 2, tanto em nível Brasil como também individualmente por PTT, fica evidente que a maior parcela dos membros interessados em *peering* consiste em provedores de acesso de abrangência local. Esse resultado era esperado e pode ser explicado pelo interesse econômico dos provedores de acesso em trocar o máximo de tráfego possível através de acordos multilaterais na infraestrutura compartilhada dos PTTs, uma vez que dessa maneira há economia decorrente da não utilização do trânsito provido pelas grandes operadoras de telecomunicações de abrangência estadual ou nacional. A iniciativa de conexão dos provedores menores aos PTTs é importante porque traz impacto direto nos preços praticados aos clientes em um cenário local de competição entre vários provedores de acesso.

A única exceção ocorreu no PTT-DF em que a presença do governo é predominante em relação às demais categorias, uma particularidade regional da capital federal do Brasil que possui como membros de destaque: o Senado Federal, a Polícia Federal, o Serviço Federal de Processamento de Dados (Serpro), a Empresa de Tecnologia da Informação da Previdência Social (Dataprev), a Telebras, entre outros.

Tabela 2. Perfil de SAs nos PTTs do Brasil

Classificação	Brasil (*)	DF	MG	RS	SP	VIX
1. Provedor de Internet	65,1% ± 20%	37,5%	55,9%	68,0%	73,1%	75,0%
1.1 Provedor de Trânsito	8,6% ± 09%	20,8%	14,7%	5,0%	5,6%	10,0%
1.2 Provedor de Acesso	56,5% ± 21%	16,7%	41,2%	63,0%	67,5%	65,0%
2. Provedor de Serviços	10,1% ± 07%	8,3%	8,8%	5,0%	12,5%	5,0%
2.1 Provedor de Conteúdo	3,2% ± 06%	0,0%	2,9%	3,0%	4,7%	0,0%
2.2 Provedor de Hospedagem	6,8% ± 05%	8,3%	5,9%	2,0%	7,8%	5,0%
3. Organização Pública	12,3% ± 21%	37,5%	20,6%	11,0%	4,4%	15,0%
3.1 Universidade Pública	1,8% ± 19%	0,0%	0,0%	2,0%	1,1%	0,0%
3.2 Governo	8,8% ± 13%	33,3%	17,6%	8,0%	2,2%	15,0%
3.3 Outros	1,8% ± 03%	4,2%	2,9%	1,0%	1,1%	0,0%
4. Organização Privada	12,6% ± 09%	16,7%	14,7%	16,0%	10,0%	5,0%
4.1 Universidade Privada	0,7% ± 03%	0,0%	2,9%	4,0%	0,0%	0,0%
4.2 Empresa Privada	10,4% ± 09%	16,7%	8,8%	10,0%	8,9%	5,0%
4.3 Outros	1,5% ± 09%	0,0%	2,9%	2,0%	1,1%	0,0%

(*) Média de todos os 26 PTTs brasileiros.

Foi observada baixa participação dos provedores de conteúdo nos PTTs em operação nas diversas regiões brasileiras, a exemplo de jornais, revistas, rádios, emissoras de televisão, etc; sendo que a maioria dos representantes da classe *provedor de conteúdo* são empresas que operam *Content Delivery/Distribution Networks* (CDN). Esse baixo resultado evidencia que ainda são poucos os provedores de conteúdo explorando os benefícios de desempenho do *peering* nos PTTs, principalmente no que diz respeito à menor latência ao diminuir o diâmetro da comunicação fim-a-fim. Por outro lado, reconhecemos como explicação plausível desse número o interesse dos provedores de conteúdo na contratação das grandes empresas de CDN para distribuição do seu conteúdo nas proximidades dos seus usuários (clientes) no panorama de um vasto território geográfico, em contraste aos PTTs que trazem benefícios apenas no panorama local.

Também é possível notar a baixa presença de empresas privadas nos PTTs, o que pode ser explicado pela motivação de redundância das empresas em implementar conexões *multi-homed* através de duas ou mais operadoras de telecomunicações que tenham alcance a toda a Internet, em contraste aos PTTs que, normalmente, têm alcance limitado à sua região geográfica.

Foi constatado que 97,72% dos SAs brasileiros conectados nos PTTs optaram por *peering* aberto com acordo de tráfego multilateral (ATM), um número bastante alto que evidencia os esforços da política pública do NIC.br no projeto PTTMetro. Apenas 2,28% optou por *peering* privado através de acordos de tráfego bilateral (ATB), sendo que praticamente a totalidade dessa porcentagem são provedores de trânsito, a destacar: Brasil Telecom, Embratel, Global Crossing, Level3, NTT, Oi, Telefonica e TIM. Essa observação é coerente porque as grandes operadoras de telecomunicações que vendem trânsito para provedores de acesso locais e regionais não têm motivação econômica para trocar tráfego abertamente com qualquer outro SA que não seja de natureza similar, uma vez que provedores menores são potenciais clientes interessados na compra de trânsito. Como exceção, as seguintes grandes operadoras optaram por ATM: Algar Telecom, Brasil Telecom (em Brasília), Cemig Telecom, Claro, CTBC, GVT, Internexa (Colombiana), NET, Nextel, Oi (em BH, Londrina, Manaus e Natal), Sercomtel e Vivo.

4.2. Graus dos Vértices

As figuras 2 e 3 mostram os graus dos vértices dos grafos referentes aos PTTs: DF, MG e VIX. Como a quantidade de vértices é muito grande para ser representada individualmente, optamos por ordenar o eixo x de maneira crescente por número de SA nos gráficos da figura 2. Dessa maneira é possível fazer uma leitura do grau médio dos vértices que equivalem a uma faixa específica de SAs, manobra que evidenciou que os SAs registrados há mais tempo (com números menores) têm maior grau médio, um resultado coerente se considerarmos que os vértices com maior grau são de operadoras de telecomunicações que têm mais adjacências pela natureza da atividade de venda de tráfego para provedores de acesso. Cada barra que representa o grau médio de uma faixa de SAs está acompanhada de um indicador do intervalo de confiança. Além dos graus médios, a figura 3 traz a distribuição dos valores exatos de graus dos nós.

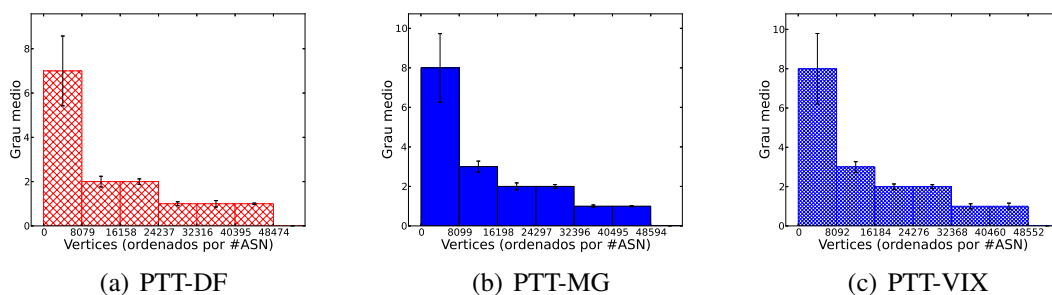


Figura 2. Grau Médio dos Grafos por Números de SAs

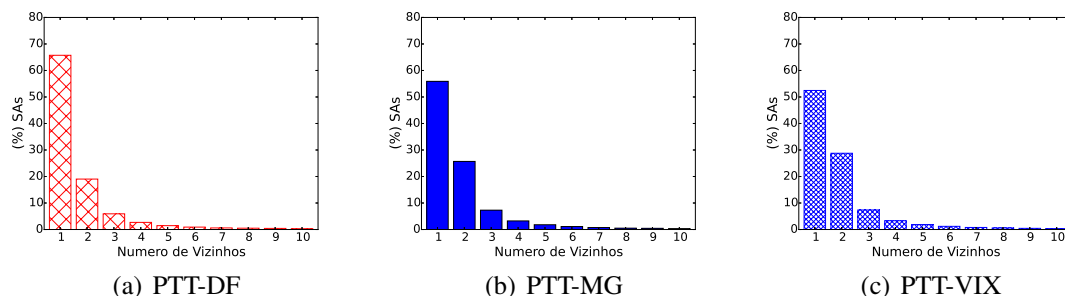


Figura 3. Distribuição dos Graus dos Grafos

4.3. Profundidade/Diâmetro

O gráfico (a) da figura 4 traz um estudo da profundidade/diâmetro dos anúncios observados na tabela BGP dos PTTs com base no atributo AS-PATH. Ao observar os números de SAs que compõem o AS-PATH é possível determinar a quantidade de saltos e quais SAs devem ser atravessados por um pacote até que a origem anunciante de um determinado prefixo seja alcançada por um membro do PTT. Em relação a essa análise destacamos que identificamos e removemos informações duplicadas que poderiam comprometer os resultados, principalmente decorrentes da aplicação da política de AS-Prepend para fins de engenharia de tráfego. Um destaque em relação aos gráficos da figura 4 é que a profundidade igual a 1 indica que o AS-PATH é composto de apenas um SA, o que deve ser interpretado como SAs diretamente conectados no PTT que estão efetivamente fazendo

anúncio dos seus próprios prefixos (baixa porcentagem). As demais rotas com profundidade maior que 1 estão sendo aprendidas pelos membros dos PTTs a partir de outros SAs, ou seja, não estão sendo diretamente anunciadas pelos seus membros adjacentes.

Tomando o PTT-VIX no gráfico (a) como referência, por exemplo, lê-se que aproximadamente 40% (eixo y) do total de rotas anunciadas na tabela BGP possuem 5 SAs e outros 40% possuem 6 SAs no atributo AS-PATH, ou seja, cerca de 80% das rotas têm profundidade de ordem 5 ou 6. Ao observar os PTTs de forma conjunta, tem-se que a profundidade média das rotas anunciadas nos PTTs varia entre 4 e 6, sendo que há maior concentração de rotas com profundidade 5. Outro detalhe relevante é que a maior profundidade nos gráficos é de ordem 7, ou seja, as origens mais distantes dos anúncios que chegam nos PTTs estão até 7 saltos de distância, sendo que cada salto representa um SA (e não um roteador). No entanto, é importante destacar que nas tabelas BGP dos PTTs existem prefixos com profundidade maior que 7 sendo anunciados, mas a quantidade desses prefixos é mínima em relação à totalidade de anúncios. Esse fato fica evidente no gráfico (b) da figura 4 que apresenta uma contagem de saltos de forma cumulativa.

Os resultados das análises sobre o grafo do PTT-SP e do PTT-PR são discrepantes e devem ser desconsiderados porque os dados extraídos a partir de acesso público à tabela BGP (via LG) sofreram uma série de filtros que comprometeram essa massa de dados. Optamos por incluir o PTT-SP na apresentação dos resultados justamente para registrar esse alerta como um dos problemas que enfrentamos durante os trabalhos de análise e para evidenciar as implicações negativas de filtros em pesquisas dessa natureza.

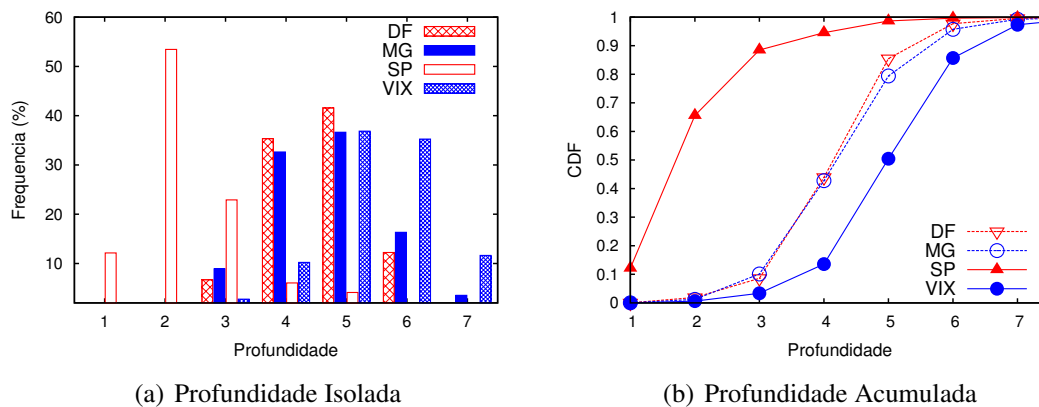


Figura 4. Profundidade dos AS-PATH

4.4. Densidade de Peering

Além do potencial natural de ampliação da troca de tráfego no ecossistema de PTTs através da participação futura de novos membros, o estudo da densidade de pares correspondente à tabela 3 evidencia que há grande dispersão nas diferentes regiões de operação dos PTTs, com potencial real de aumento na troca de tráfego apenas entre os membros já existentes. A densidade de pares consiste na razão entre a quantidade de conexões (peering) existentes em um PTT com sua combinação de pares possíveis.

4.5. Engenharia de Tráfego Baseada em AS-Prepend

À medida em que os prefixos anunciados atravessam domínios BGP, o número do SA anterior é inserido à lista de SAs que foram percorridos, com o objetivo de descartar

Tabela 3. Densidade de Pares nos PTTs

Descrição da Métrica	Brasil (*)	DF	MG	RS	VIX
Peers Existentes (Conexões)	126	57	79	1.952	71
Combinação de Peers Possíveis	285	276	231	3.081	153
Densidade (%)	44,2% ± 23%	20,7%	34,2%	63,4%	46,4%

(*) Média dos 24 PTTs brasileiros sem filtro, ou seja, excluindo PTT-PR e PTT-SP.

conteúdo que tenha seu próprio SA no AS-PATH para evitar *loops*. No comportamento padrão do BGP as menores listas de AS-PATH serão preferíveis, ou seja, aqueles caminhos com menor profundidade/diâmetro, o que torna o atributo AS-PATH importante para selecionar a melhor rota quando existem múltiplas rotas para um mesmo destino. O atributo AS-PATH é frequentemente manipulado por um SA para adicionar (*prepend*) mais de uma vez seu próprio número de SA e, portanto, tornar menos atrativa a alcançabilidade de um prefixo através dele mesmo.

Os números da tabela 4 reforçam que AS-Prepend é frequentemente praticado na Internet, conforme pode ser observado nos estudos da quantidade total de rotas que sofreram AS-Prepend nos *Paths BGP* dos PTTs e também de SAs que praticam essa técnica para fins de políticas de roteamento. Cabe observar que o segundo bloco da tabela diz respeito a todos os SAs na Internet que foram observados a partir dos grafos dos PTTs, sejam eles membros ou não. No terceiro bloco da tabela são trazidos os números referentes apenas aos membros diretamente conectados nos PTTs, sendo importante observar que nem todos os membros de um PTT fazem anúncios.

Uma observação relevante é que o número de rotas nas tabelas BGP é um conceito diferente do número de prefixos. Sabe-se que atualmente a tabela BGP completa possui pouco mais de 512.000 prefixos, no entanto é comum as tabelas BGP dos PTTs possuírem milhares de linhas porque podem ter múltiplas rotas apontando para um mesmo prefixo.

Tabela 4. Ocorrência de AS-Prepend na Visão dos PTTs

Descrição da Métrica	Brasil (*)	DF	MG	RS	VIX
Quantidade de Rotas	932.989	559.159	434.264	1.947.453	2.663.751
Quantidade de Rotas com AS-Prepend	295.909	127.184	245.129	1.710.070	623.965
AS-Prepend X Rotas (%)	30,8% ± 22%	22,7%	56,4%	87,8%	23,4%
Quantidade de SAs	43.333	47.176	46.939	48.351	47.474
Quantidade de SAs com AS-Prepend	7.305	6.206	8.629	10.803	9.124
AS-Prepend X SAs (%)	16,1% ± 4%	13,2%	18,4%	22,3%	19,2%
Quantidade de Membros (Anunciantes)	18	24	22	79	18
Quantidade de Membros c/ AS-Prepend	6	7	6	36	5
AS-Prepend X Membros (%)	22,5% ± 19%	29,2%	27,3%	45,6%	27,8%

(*) Média dos 24 PTTs brasileiros sem filtro, ou seja, excluindo PTT-PR e PTT-SP.

4.6. Comunidades k-Clique

Uma comunidade é composta por um conjunto de SAs (vértices), sejam eles membros diretamente conectados em um PTT ou mesmo externos. Os métodos de extração de comunidades da Internet que se destacam na literatura são *k-clique*, *k-dense* e *k-core* [Orsini et al. 2013]. A principal característica em comum desses métodos é detectar conjuntos de SAs densos no que diz respeito à conectividade.

O método de comunidades k-clique foi escolhido porque possui característica peculiar de fornecer comunidades sobrepostas [Palla et al. 2005], o que permite separar comunidades que tenham certa quantidade de nós em comum. Um k-clique é um subgrafo completo com k nós. Dois k-cliques são adjacentes se possuírem $k - 1$ nós em comum. Uma comunidade k-clique, ou *comunidade(k)*, é o conjunto maximal formado pela união de todos os k-cliques adjacentes. Outra propriedade é que em cada comunidade k-clique existe uma e apenas uma comunidade (k-1)-clique tal que *comunidade(k)* é um subgrafo de *comunidade(k-1)* [Gregori et al. 2011], ou seja, com menores valores de k é possível identificar mais comunidades do que seria identificado com maiores valores de k .

Através desse estudo foi possível encontrar comunidades e caracterizar seus membros de maneira a ultrapassar a perspectiva localizada dos PTTs, fornecendo uma visão da Internet propriamente dita. Para valores menores de k , principalmente 3 ou 4, tendo como referência a visão dos PTTs estudados, a quantidade de comunidades encontradas na Internet é similar (figura 5(a)), sendo que essas comunidades possuem alta densidade de conexão entre seus membros (figura 5(b)). Igualmente à densidade previamente apresentada no contexto específico dos PTTs (sub-seção 4.4), definimos a densidade desse estudo como a média das densidades individuais das comunidades k-clique.

Considerando a caracterização previamente apresentada na sub-seção 4.1, uma observação interessante no contexto do ecossistema dos PTTs foi a identificação de comunidades k_{max} , cujos membros têm perfil similar. Definimos uma comunidade k_{max} como um conjunto não vazio de SAs que forma uma comunidade k-clique com o valor máximo de k . Existem poucas comunidades k_{max} que possuem SAs com identificadores (ASN) de 3 ou 4 dígitos em sua grande maioria. Tais SAs são predominantemente provedores de trânsito *tier-1*, com poucos casos de presença de grandes empresas como Facebook, Microsoft, Amazon, etc. Em comunidades com menores valores de k existem mais SAs provedores de acesso (com ASNs de 5 dígitos). A figura 5(c) mostra a quantidade de SAs das maiores comunidades encontradas, além de permitir inferir os valores de k_{max} , observando que o PTT-DF não aparece no gráfico porque seu k_{max} é apenas 6.

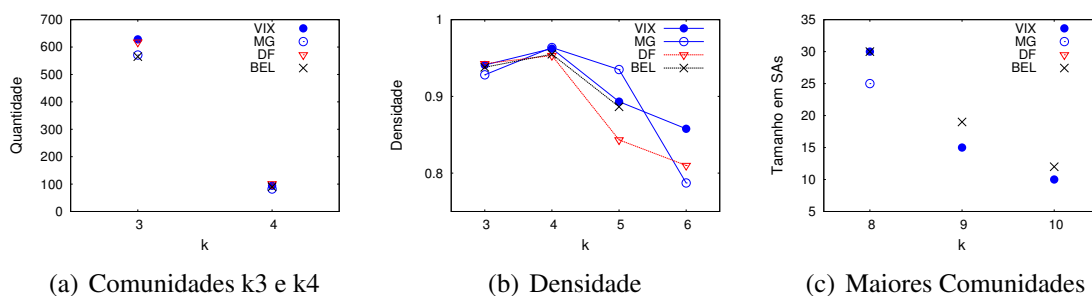


Figura 5. Comunidades k-Clique

5. Limitações

Igualmente importante ao esforço de análise apresentado nesse artigo é o reconhecimento de algumas limitações de pesquisas intencionadas em construir uma topologia da Internet em nível de SAs [Haddadi and Bonaventure 2013]. Embora tenhamos confiança nos resultados da nossa metodologia de análise do *dataset* que estamos compartilhando, é importante destacar que as informações extraídas dos servidores públicos não representam

a totalidade da troca de tráfego que realmente ocorre entre os diversos membros dos SAs brasileiros, mas apenas uma fração equivalente a tudo aquilo que pode ser publicamente observado. Em termos de *peering* nos PTTs, somente os acordos de tráfego multilaterais (ATM) são espelhados nos servidores públicos e podem ser visualizados a partir dos LGs, sendo que os acordos de tráfego bilaterais (ATB) não podem ser visualizados porque são diretamente estabelecidos entre dois SAs de maneira privada.

Outro detalhe relevante é que os LGs públicos deveriam espelhar todas as rotas presentes no(s) RS(s) dos PTTs, no entanto é impossível garantir que as observações a partir do acesso público não tenham sofrido filtragem prévia por alguma política administrativa na gestão dos PTTs. Por exemplo, durante nossas análises ficou evidente que as tabelas BGP extraídas do PTT-SP e do PTT-PR não estavam completas ao compará-las com as tabelas dos demais PTTs, informação que foi confirmada pelo NIC.br com a justificativa de problemas de escalabilidade. Essa limitação explica a maior concentração de rotas com menor profundidade e a baixa porcentagem de densidade resultante da análise do PTT-SP, enquanto que nos demais PTTs há maior concentração de rotas com maior profundidade e maior porcentagem de densidade.

Uma das informações constantes no *dataset* que compartilhamos são os resumos das *communities* BGP configuradas nos PTTs, no entanto não pudemos utilizar esses dados para inferir *links* ocultos seguindo a metodologia proposta em [Giotsas et al. 2013] porque fomos informados pelo NIC.br de que não existe um sistema de codificação padronizado para utilização das *communities* no PTTMetro brasileiro.

6. Trabalhos Futuros

Uma primeira extensão importante da nossa metodologia é a construção de uma análise temporal a partir de amostras de dados de períodos futuros, fornecendo diferentes *snapshots* que permitam compreender a evolução da dinâmica do ecossistema de PTTs no Brasil e construir um modelo de predição. Também pretendemos expandir nosso estudo no sentido de realizar uma análise dos tamanhos dos blocos de prefixos IPv4 e IPv6 constantes nas tabelas BGP, o que permitirá comparar os dois protocolos e verificar qual perfil de membro tende a anunciar qual tamanho de bloco.

Estamos trabalhando em análises mais aprofundadas sobre as comunidades clique porque enxergamos a possibilidade de inferir novos resultados, principalmente no sentido de caracterização dos SAs e de compreensão da Internet. Pretendemos expandir nossa visão das comunidades na Internet a partir da extração de dados de servidores de rotas disponibilizados publicamente pelas grandes operadoras de telecomunicações. Esse conjunto de dados será utilizado para subsidiar pesquisas relacionadas a *SDN Peering* que estamos desenvolvendo na perspectiva de que a combinação de SDN com PTTs viabilize novas aplicações que não são possíveis atualmente [Gupta et al. 2014], como por exemplo o *peering* baseado em conteúdo ou aplicação e novas maneiras de aplicar políticas entre domínios para fins de engenharia de tráfego.

7. Conclusão

Este artigo apresentou um estudo que acreditamos ter sido o primeiro esforço de compreensão do ecossistema de PTTs públicos em operação no Brasil, compilando informações

coletadas a partir de servidores públicos que nos permitiram reproduzir os grafos de conectividade dos PTTs brasileiros para compreensão da sua anatomia, oportunidade em que identificamos diversos aspectos da estrutura desse ecossistema.

Agradecimentos

Agradecemos ao Centro de Inovação da Ericsson Telecomunicações S/A do Brasil pelo suporte financeiro. Também ao NIC.br pela colaboração, especialmente aos colegas Antonio Moreiras (CEPTRO.br) e Antonio Galvão de Rezende Filho (PTT.br).

Referências

- Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S., and Willinger, W. (2012). *Anatomy of a Large European IXP*. SIGCOMM'12, August 13-17, Helsinki, Finland.
- Chatzis, N., Smaragdakis, G., Feldmann, A., and Willinger, W. (2013). *There is More to IXPs than Meets the Eye*. CCR ACM SIGCOMM. Volume 43, Number 5. October.
- Chen, K. and et al (2009). *Where the Sidewalk Ends, Extending the Internet AS Graph Using Traceroutes from P2P Users*. CoNEXT'09, December 01-04, Rome, Italy.
- Giotsas, V., Luckie, M., Huffaker, B., and Claffy, K. (2014). *Inferring Complex AS Relationships*. IMC'14. November 6-7. Vancouver, BC, Canada.
- Giotsas, V., Zhou, S., Luckie, M., and Claffy, K. (2013). *Inferring Multilateral Peering*. CoNEXT'13, December 9-12. Santa Barbara, California, USA.
- Gregori, E., Lenzini, L., and Orsini, C. (2011). *k-clique Communities in the Internet AS-level Topology Graph*. 31st ICDCSW. IEEE.
- Gupta, A., Vanbever, L., Shabaz, M., Donovan, S. P., Schlinker, B., Feamster, N., et al. (2014). *SDX - A Software Defined Internet Exchange*. SIGCOMM'14. Chicago, USA.
- Haddadi, H. and Bonaventure, O. (2013). *Recent Advances in Networking. Chapter 1: Internet Topology Research Redux*. ACM SIGCOMM eBook. Volume 1. August.
- Labovitz, C., Iekel-Johnson, S., McPherson, D., Oberheide, J., and Jahanian, F. (2010). *Internet Inter-Domain Traffic*. SIGCOMM'10, Aug 30 - Sep 3, New Delhi, India.
- Lodhi, A., Larson, N., Dhamdhere, A., Dovrolis, C., and Claffy, K. (2014). *Using PeeringDB to Understand the Peering Ecosystem*. CCR ACM SIGCOMM. April.
- Luckie, M., Huffaker, B., Dhamdhere, A., Giotsas, V., and Claffy, K. (2013). *AS Relationships, Customer Cones, and Validation*. IMC'13. Barcelona, Spain.
- Norton, W. B. (2014). *The Internet Peering Playbook: Connecting to the Core of the Internet*. DrPeering Press. 2014 Edition. USA.
- Orsini, C., Gregori, E., Lenzini, L., and Krioukov, D. (2013). *Evolution of the Internet k-Dense Structure*. IEEE/ACM Transactions on Networking. September 16.
- Palla et al. (2005). *Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society*. Nature. Volume 435. Number 7043.
- Restrepo, J. C. C. and Stanojevic, R. (2012). *A History of an Internet eXchange Point*. CCR ACM SIGCOMM. Volume 42, Number 2, April.
- Schlinker, B., Zarifis, K., Cunha, I., Feamster, N., and Katz-Bassett, E. (2014). *PEERING: An AS for US*. HotNets'14. October 27-28. Los Angeles, California, USA.