



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Ausarbeitung Anwendung 1
WiSe 2010/2011
Andreas Krohn
Erweiterung des Routing-Atlas

Inhaltsverzeichnis

1 Einführung	3
2 Routing-Atlas	3
2.1 Identifikation des „deutschen Internets“	4
2.2 Bildung des AS Graphen	4
2.3 Visualisierung	5
3 Zielsetzung	6
4 Konzept	7
4.1 Erstellung der next hop matrix nach Winter	7
4.2 Adaption für den Routing-Atlas	9
5 Herausforderungen	10
6 Ausblick	10
7 Zusammenfassung	11
Literatur	12

1 Einführung

Die Bedeutung des Internets nimmt in den letzten Jahren stetig zu. Kommunikation, Verwaltung und Abwicklung von Geschäften verlagern sich zunehmend ins Internet. Gleichzeitig wachsen die Bestrebungen seitens der Politik und Unternehmen Einfluss auszuüben - sei es zur Durchsetzung von Gesetzen oder zur Gewinnmaximierung. Angesichts dieser Tatsache ist es interessant zu analysieren, welche Akteure die Daten beispielsweise auf dem Weg zwischen Geschäftspartnern oder zwischen einer Behörde und Bürgern traversieren. Um dieses Wissen zu erlangen, sind Topologieanalysen nötig, wie sie das Routing-Atlas Projekt durchführt.

In dieser Ausarbeitung wird der Routing-Atlas sowie die geplanten Änderungen vorgestellt.

In Abschnitt 2 wird zunächst auf die bereits vorhandenen Schritte und Datenquellen zur Erstellung des Routing-Atlas eingegangen. Abschnitt 3 umreißt die Zielsetzung, deren Hauptbestandteil - Ersatz einer externen, nicht mehr aktualisierten Datenquelle - in Abschnitt 4 erläutert wird. Abschnitt 5 zeigt, welche Herausforderungen erwartet werden. Den Abschluss bilden der Ausblick auf weitere Ausbaumöglichkeiten des Routing-Atlas in Abschnitt 6 und die Zusammenfassung in Abschnitt 7

2 Routing-Atlas

Der Routing-Atlas ist ein Projekt der inet AG in Zusammenarbeit mit dem Bundesamt für Sicherheit in der Informationstechnik (BSI), vgl. (1; 2; 3). Angesichts der wachsenden Bedeutung des Internet als kritische Infrastruktur für Wirtschaft, Verwaltung und Bildung und der zunehmenden Regulierung dieses Mediums stellt sich die Frage welche Akteure und Jurisdiktionen Einfluss nehmen können. Ziel des Projektes ist es landesspezifische Teile des Internet und Abhängigkeiten zwischen Ländern zu identifizieren, zu klassifizieren und zu visualisieren. Damit kann zum Beispiel abgeschätzt werden:

- Zu welchen Akteuren und Ländern bestehen Abhängigkeiten?
- Wie stark ist die landesinterne Kommunikation in sich geschlossen?
- Wie störfähig ist die landesinterne Kommunikation?
- Wie leicht lässt sich nationaler Traffic umleiten?

Mit ähnlichen Fragenstellungen haben sich bereits Karlin et al. (4) beschäftigt. Sie haben hierbei die Zuordnung zu Ländern auf der Basis teils bereits aggregierter IP-Präfixe statt der im Routing-Atlas verwendeten IP-Adressblöcken vorgenommen.

Im Folgenden wird auf die Schritte zur Erstellung des Routing-Atlas eingegangen.

2.1 Identifikation des „deutschen Internets“

Die Organisation des Internet berücksichtigt politische oder geographische Grenzen nicht. IP-Adressen, IP-Adressblöcke, IP-Präfixe und Nummern Autonomer Systeme enthalten für sich genommen keine Metainformationen, aus denen auf die Zugehörigkeit zu einem Land oder einer Organisation geschlossen werden kann. Zur Lösung dieses Problems werden Informationen aus administrativen Datenbanken der Regional Internet Registries (RIR) genutzt. Die für Europa zuständige RIR ist das Réseau IP Européens Network Coordination Centre (RIPE NCC) (5). Um möglichst viele Teilnehmer des „deutschen Internets“ zu identifizieren, werden zunächst IP-Adressblöcke - die kleinsten in der Datenbank vertretenen Entitäten - betrachtet.

Das RIPE NCC vergibt für Europa IP-Adressblöcke und verwaltet diese in einer (weitesgehend) öffentlich zugänglichen Datenbank. Diese Datenbank wird nach Adressblöcken (`inetnum`-Objekte) durchsucht, deren `country`-Attribut 'DE' oder 'EU' ist. Zu den 'EU' Adressblöcken wird in weiteren Attributen nach Einträgen gesucht, die eine Einordnung als „deutscher“ Adressblock zulassen. Beginnt beispielsweise die Telefonnummer einer für den Adressblock verantwortlichen Kontaktperson oder Organisation mit '+49', wird der Adressblock aufgenommen. Aus datenschutzrechtlichen Gründen limitiert die RIPE den Zugriff auf personenbezogenen Daten auf ein Kontingent von einigen Hundert Anfragen pro Tag, sodass einmal abgefragte Informationen zwischengespeichert werden sollten und die Frequenz der Anfragen an die RIPE DB gering gehalten werden muss.

Im nächsten Schritt werden die gefundenen Adressblöcke zu routbaren Präfixen zusammengefasst. Dazu werden `route`-Objekte in der RIPE DB gesucht, deren Präfixe die identifizierten Adressblöcke beinhalten.

Die hierbei gefundenen `route`-Objekte haben idR. ein `origin`-Attribut, dass auf den Datensatz des Autonomes Systems (`aut-num`-Objekt) verweist. Fehlt dem Präfix eine Zuordnung zu einem AS, wird eine Auflösung beim Team Cymru und - schlägt auch dies fehl - beim RIPE RIS-Kollektor RRC 12 (DE-CIX in Frankfurt) versucht (6; 7).

Durch die Nutzung mehrerer Datenquellen bleiben nur wenige Präfixe ($\approx 2\%$) unaufgelöst.

2.2 Bildung des AS Graphen

Am NEC-Lab wurde im Rahmen eines europäischen Forschungsprojektes ein Modell der Routing-Topologie des Internets auf Ebene der Autonomes Systeme erstellt (8). Hierfür wurden die vom Internet Research Lab der UCLA gesammelten und aggregierten Daten genutzt (9). Ein Ergebnis dieses Modells ist die next hop matrix. Sie beinhaltet für die Wege von einem AS zu einem anderen jeweils das nächste AS, über das ein Paket geroutet werden würde. Auf die Generierung der next hop matrix wird in Abschnitt 4.1 eingegangen.

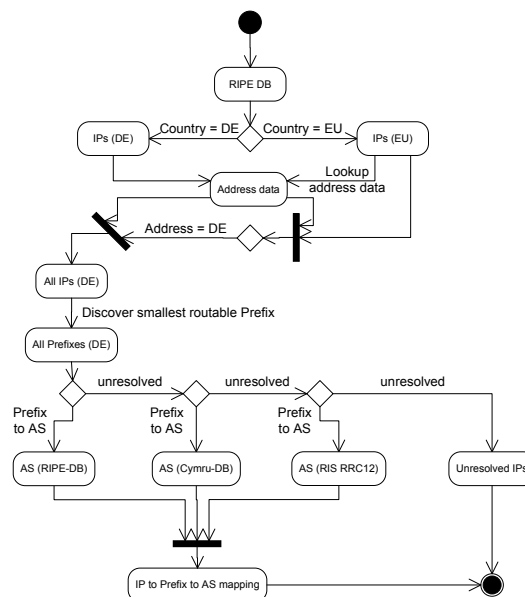


Abbildung 1: Toolchain: IP-Adressblöcke zu Autonomen Systemen

Der Routing-Atlas verwendet die next hop matrix, um einen Graphen aller im vorherigen Schritt identifizierten Autonomen Systeme zu bilden. Dieser Graph umfasst neben den „deutschen“ ASen auch jene, die für die Verbindung sorgen und im Ausland liegen - die Transit-ASen. Analysiert man nun diesen Graphen, so lassen sich Antworten auf die ursprünglichen Fragen nach der Abhängigkeit und Geschlossenheit landesinterner Kommunikation vom Ausland finden.

2.3 Visualisierung

Der im vorherigen Schritt erstellte Graph wird graphisch dargestellt. Da der Gesamtgraph in Auflösungen unterhalb eines wandfüllenden Ausdrucks unübersichtlich ist, werden hierbei hauptsächlich gefilterte Subgraphen dargestellt. Um weiterhin die Bedeutung einzelner Autonomer Systeme besser zu veranschaulichen, wird die dargestellte Einfärbung, Größe oder Position eines AS nach einer Klassifikation festgelegt. Aspekte nach denen ein AS klassifiziert wird sind

- die Anzahl enthaltener IP-Präfixe oder IP-Adressblöcke,
- die hierarchische Einordnung in die Routing-Topologie des Internets,
- die Branchenzugehörigkeit.

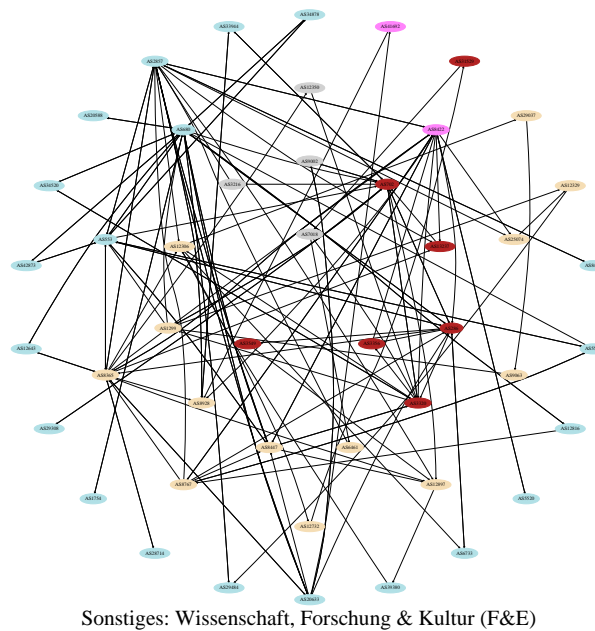


Abbildung 2: Hierarchisches Kreismodell

Abbildung 2 zeigt die Autonomen Systeme der Branche 'Wissenschaft, Forschung & Kultur (F&E)' in einem hierarchischen Kreisdiagramm. Die Position eines AS in diesem Diagramm bedeutet, dass es in seiner topologischen Bedeutung (von innen nach außen) Tier 1, Large ISP, Small ISP und Stub zuzuordnen ist.

3 Zielsetzung

Ziel des Autors ist die Erweiterung des Routing-Atlas. Die Erweiterungen lassen sich in die akut nötige Schaffung eines Ersatz für die next hop matrix des NEC-Lab und weniger dringende, dennoch interessante und erstrebenswerte Weiterentwicklungen unterteilen. Das NEC-Lab Projekt, das die in Abschnitt 2.2 vorgestellte next hop matrix erstellt, wurde zum Ende des Jahres 2009 eingestellt. Daraus folgt, dass eine wichtige Datengrundlage des Routing-Atlas seither altert. Die Topologie des Internet ist stetigen Änderungen unterworfen, die so vom Routing-Atlas nicht erfasst wird.

Primäre Zielsetzung der Erweiterungen am Routing-Atlas ist es einen kontinuierlich aktualisierten Ersatz für die next hop matrix des NEC-Lab zu schaffen. Ein Konzept hierfür wird in Kapitel 4 vorgestellt.

Daneben sind weitere Entwicklungen geplant:

Online-Aktualisierung Um Veränderungen der Topologie und deren Auswirkungen besser beobachten zu können, sollte der Routing-Atlas automatisiert regelmäßig neu generiert und die Ergebnisse - in einem zu entwerfenden interaktiven Werkzeug - dargestellt werden.

Weitere Länder Die Anwendung der Verfahren für die Generierung des deutschen Routing-Atlas auf ein anderes europäisches Land erfordert Anpassungen und ermöglicht vergleichende Analysen der Topologien verschiedener Länder.

IPv6 Die „Portierung“ des Routing-Atlas in die IPv6-Welt ist aufgrund der Aktualität dieses Themenbereichs und der hier stattfindenden Entwicklung und Bewegung interessant. Hier stellt sich die Herausforderung, dass die vom IRL der UCLA bereitgestellten Daten derzeit nur IPv4 berücksichtigen.

4 Konzept

Einen Ansatz zur Erstellung einer next hop matrix der AS Topologie wurde bereits von Winter (8) vorgestellt. Im Folgenden wird auf diesen Ansatz eingegangen und anschließend gezeigt, welche Anpassungen vorgenommen werden sollen.

4.1 Erstellung der next hop matrix nach Winter

Basierend auf den vom Topology Project des Internet Research Lab an der UCLA gesammelten Daten (9) erstellt Winter ein Modell, das die Topologie der Autonomen Systeme darstellt. Dabei wird auf die Berücksichtigung des real stattfindenden präfixbasierten Routings verzichtet und nur AS Pfade modelliert. Zum einen würde die Berücksichtigung der Präfixe den Rechenaufwand erheblich steigern, zum anderen haben Analysen von Winter gezeigt, dass AS Pfade zu ca. 75% nicht divergieren, d.h. ASe werden größtenteils unabhängig von den Präfixen über eine Route erreicht. Die geringe Verfälschung der Ergebnisse ist also gegenüber dem Gewinn einer geringeren Zeitkomplexität zu rechtfertigen.

Die Datensammlung der UCLA beinhaltet aggregiert alle in Routingtabellen und BGP-Updates beobachteten Routen - auch solche, die aktuell nicht mehr genutzt werden. Bereitgestellt werden diese Daten in Form von mehreren unterschiedlich vorverarbeiteten Dateien. Interessant für die Konstruktion eines AS Graphen ist das „Link file format“. Die Datei enthält zeilenweise beobachtete Verbindungen zwischen zwei Autonomen Systemen mit Quell- und Ziel-AS Nummer, Zeitstempel der ersten und letzten Beobachtung, die Position dieser Verbindung innerhalb

der beobachteten Route sowie die Rohdaten aus denen die Verbindung extrahiert wurde, vgl. [Abbildung 3](#)

```
AS0    AS1    first seen    last seen    pos    last bgp update
14288  25642  0             1297152006  1     TABLE_DUMP2|1297152006|B|8..
8468   559    1226590780   1297151998  1     TABLE_DUMP2|1297151998|B|1..
1853   9002   1214921279   1297152004  0     TABLE_DUMP2|1297152004|B|1..
3130   21143  1235631042   1240124919  0     BGP4MP|1240124919|A|147.28..
...

```

Abbildung 3: Ausschnitt einer Datei im „Link file format“

Zhang et al. (10) haben gezeigt, dass Routen, für die für die Dauer eines Monats kein erneutes BGP-Announcement beobachtet wird, als nicht mehr vorhanden betrachtet werden können. Auf diese Annahme gestützt werden die AS Links vor der Modellierung des AS Graphen nach dem letzten Beobachtungszeitpunkt gefiltert und nur die im letzten Monat beobachteten Routen verwendet.

Aus den verbleibenden AS Links werden mittels Anwendung des Floyd-Warshall-Algorithmus die kürzesten Pfade zwischen allen AS Paaren berechnet - die ungewichtete shortest path und next hop matrix.

BGP und damit das Routing im Internet arbeitet jedoch nicht alleine basierend auf Pfadlängen. Es berücksichtigt Policies, die basierend auf Geschäftsbeziehungen gesetzt werden und regeln, wie „gerne“ möglichen next hops verwendet werden. Um diese Tatsache in das Modell einfließen zu lassen, werden die gerichteten AS Links mit Gewichten belegt. Diese Gewichte basieren auf einer weiteren, vom IRL der UCLA bereitgestellten Datei - der „weights“ Datei. In der Datei sind zeilenweise zu einem Link unter anderem Quell- und Ziel-AS Nummer sowie die Anzahl der Beobachtungen in Routingtabellen enthalten.

Das Gewicht eines Links für die Berechnung der gewichteten shortest path und next hop matrix wird aus Anzahl der Beobachtungen dieses Links in Routingtabellen errechnet. Links, die nie in Routingtabellen auftauchen, werden mit einem Gewicht von 1000 belegt. Anderen Links wird ein Gewicht zwischen 1 und 100 zugewiesen. Wobei 1 bedeutet, dass alle ausgehenden Pfade des Quell-AS durch diesen Link verlaufen, bei einem Wert von 50 die Hälfte der Verbindungen und bei 100 weniger als 1% der Verbindungen - also: je höher das Kantengewicht, desto unbedeutender/seltener beobachtet ist ein Link. Die Vorbelegung mit 1000 ist nötig, da nur eine begrenzte Anzahl an Routingtabellen in die Datenbasis einfließen, Routingtabellen immer nur eine lokale Sicht auf das globale Routing bieten und somit nicht alle verwendeten Links in den vorhandenen Routingtabellen auftauchen, obwohl sie verwendet werden.

Auf die gewichteten AS Links wird der Floyd-Warshall-Algorithmus angewendet und so die gewichtete shortest path und next hop matrix erstellt.

Das NEC-Lab hatte für diese Berechnungen einen NEC SX-8 Vector Computer zur Verfügung, was die Anwendung gleichartiger Operationen auf eine große Datenmenge stark beschleunigt. Trotz dieser spezialisierten Hardware und der Beschränkung auf AS Links dauerte die Berechnung der globalen shortest path und next hop matrix mehrere Stunden - Winters Arbeit (8) trägt die 24 Stunden bereits im Titel.

4.2 Adaption für den Routing-Atlas

Für den Routing-Atlas ist gegenüber dem von Winter gewählten Ansatz nur der Teil des globalen AS Graphen interessant, der die als deutsch identifizierten Autonomen Systeme untereinander verbindet. Diese Reduzierung des Problems berechtigt zu der Hoffnung, die Berechnung einer next hop matrix ohne Vektor Computer auf einem Mehrprozessorsystem mit Allzweckprozessor in akzeptabler Zeit zu absolvieren.

Die für die Berechnung dieser Matrix nötigen Links können durch die Betrachtung der näheren Umgebung der „deutschen ASe“ gefunden werden. Analysen von Schmidt et al. haben gezeigt, dass die Pfadlängen zwischen besagten ASen typischerweise einer „Gaußverteilung mit Mittelwerten zwischen 3.0 und 3.5 sowie Streuung von ≈ 2 “ folgen (3, S. 15). Ausgehend von den identifizierten ASen werden solange Links zu angrenzenden ASen hinzugenommen, bis alle identifizierten ASe verbunden sind. Dieses Verfahren sollte aufgrund der o.g. Verteilung der Pfadlängen schnell konvergieren.

Aus der dabei erhaltenen Menge von Links wird dann nach dem Vorbild Winters eine shortest path und next hop matrix erstellt. Ob hierbei auf der von Winter verwendeten Floyd-Warshall-Algorithmus oder andere wie Dijkstra oder A* das günstigste Laufzeitverhalten liefern ist zu evaluieren. Die Wahl der Kantengewichte wird zunächst von Winter übernommen. Bei der Implementierung ist darauf zu achten die Datenstrukturen und Abläufe so zu wählen, dass die 24 Kerne des zur Verfügung stehenden Systems gut ausgelastet werden.

Die errechnete shortest path und next hop matrix muss validiert werden. Hierfür sind mehrere Ansätze denkbar. Zum einen könnten die Berechnungen auf den weiterhin verfügbaren Daten des IRL aus 2009 durchgeführt werden und die hierbei entstandenen Ergebnisse mit denen von Winter verglichen werden. Zum anderen könnten die auf aktuellen Daten errechneten Routen mit den an Route-Kollektoren beobachteten Routen verglichen werden. Dabei zu Tage tretende Abweichungen sind zu analysieren. Soweit sie nicht auf die eingeschränkte Sicht der Punkte, an denen die Referenzdaten erhoben werden, zurückzuführen sind, muss die Berechnung zur Behebung dieser Abweichungen korrigiert werden. Ergibt die Validierung große Abweichungen vom erwarteten Resultat, sollte die Gewichtung der Links neu überdacht werden. Hierbei gäbe es die Möglichkeit statt der Beobachtungen der Links in Routingtabellen eine Auswertung der BGP Updates zur Gewichtung der Links zu nutzen.

5 Herausforderungen

Die offensichtliche Herausforderung ist die, trotz der im vorigen Abschnitt beschriebenen Verringerung der zu berücksichtigenden Linkanzahl, hohe algorithmische Komplexität bei der Berechnung der next hop matrix - für Floyd-Warshall beispielsweise $O(n^3)$. Umso mehr ist bei der Umsetzung der Verfahren auf gute Skalierbarkeit zu achten. Sollte auf der verfügbaren Hardware die erforderliche Rechenzeit jenseits akzeptabler Werte liegen, wäre es so möglich die Berechnungen auf weitere Systeme zu verteilen.

Die begrenzt zur Verfügung stehenden Rechenleistung und das Ziel der Reduktion der Rechenzeit zwingen zu Kompromissen bei der Modellierung. Diese sind zum einen der Verzicht auf die Berücksichtigung der Präfixe bei der Generierung der shortest path und next hop matrix und zum anderen die Wahl der Gewichte für die Links. Beide erfolgen mit dem Ziel das Problem handhabbar zu machen und laut Winters Analysen ohne zu großen Einfluss auf die Aussagekraft des Ergebnisses. Vor diesem Hintergrund ist zu zeigen, dass die vorgeschlagene Verringerung der für die Modellgenerierung betrachteten Links ebenso keine (oder nur eine sehr geringe) Verfälschung der Ergebnisse verursacht.

Topologieanalysen des Internets bergen die Schwierigkeit, dass es keine Instanz gibt, die die tatsächliche globale Topologie kennt. Die Struktur der zu beobachtenden Pfade hängt stark vom Punkt der Beobachtung ab. Es gibt eine Menge an Projekten, die Daten sammeln und so Hinweise auf die Topologie geben. Hier ist die Interpretierung von Übereinstimmung und Abweichung zwischen modellierter und beobachteter Topologie die Herausforderung.

6 Ausblick

Mit einer neu berechneten shortest path und next hop matrix wird die Aktualität der Datenbasis des Routing-Atlas vorerst wiederhergestellt. Darauf basierend sollen die Bestandteile des Routing-Atlas soweit automatisiert werden, dass sie regelmäßig ausgeführt und Veränderungen über die Zeit dargestellt werden können. Diese Darstellungen sollen in Form eines Online-Dienstes bereitgestellt werden.

Um vergleichende Analysen zwischen verschiedenen Ländern zu ermöglichen, sollen die Methoden des Routing-Atlas auf andere (europäische) Länder angewendet werden. Der Fokus liegt hierbei auf Europa, weil dies zum einen die nähere Umgebung betrifft und zum anderen die Qualität der Daten der europäischen RIR (RIPE NCC) gegenüber denen anderer RIRs gut ist.

Weiterhin besteht der Plan, die momentan auf das IPv4-Protokoll bezogenen Methoden des Routing-Atlas auf die Anwendung auf IPv6 zu übertragen. Hierbei ist der Flaschenhals derzeit

die auf IPv4 begrenzte Datensammlung des IRL. Routekollektoren und RIPE DB unterstützen IPv6 bereits. Die Daten des IRL sind eine wertvolle Quelle für im Rahmen des Routing-Atlas durchgeführten Topologieanalysen. Es wäre also vor der Portierung des Routing-Atlas durchaus sinnvoll die (hoffentlich baldige) Integration der Beobachtung von IPv6 seitens des IRL abzuwarten. Der Aufwand auch für diese Sammlung einen Ersatz zu schaffen dürfte relativ hoch sein.

7 Zusammenfassung

In dieser Ausarbeitung wurde der Routing-Atlas in seiner derzeitigen Form vorgestellt. Weiterhin wurde Winters Konzept zur Generierung von shortest path und next hop matrix, sowie die geplanten Anpassungen um diese Generierung in die Toolchain des Routing-Atlas zu integrieren, gezeigt.

Die dabei erwarteten Herausforderungen und mögliche Abhilfen wurden aufgezeigt.

Der Ausblick zeigt weitere Betätigungsfelder auf, die nach der Lösung der shortest path und next hop matrix Problematik in Angriff genommen werden sollen.

Literatur

- [1] M. Wählisch, T. Schmidt, S. Meiling, M. de Brün, and T. Häberlen, "Towards a nation-centric understanding of the internet," in *Proceedings of the ACM CoNEXT Student Workshop*. ACM, 2010, p. 7.
- [2] M. Wählisch, S. Meiling, and T. Schmidt, "A framework for nation-centric classification and observation of the internet," in *Proceedings of the ACM CoNEXT Student Workshop*. ACM, 2010, p. 15.
- [3] T. C. Schmidt, M. Wählisch, M. de Brün, and T. Häberlen, "Ein Routing-Atlas für die strukturelle und visuelle Exposition des deutschen Internets," 2010.
- [4] J. Karlin, S. Forrest, and J. Rexford, "Nation-State Routing: Censorship, Wiretapping, and BGP," Mar. 2009. [Online]. Available: <http://arxiv.org/abs/0903.3218v1>
- [5] Ripe network coordination centre. [Online]. Available: <http://www.ripe.net/>
- [6] Ris raw data. [Online]. Available: <http://www.ripe.net/data-tools/stats/ris/ris-raw-data>
- [7] Team cymru. [Online]. Available: <http://www.team-cymru.org/>
- [8] R. Winter, "Modeling the internet routing topology - in less than 24h." in *PADS*. IEEE Computer Society, 2009, pp. 72–79. [Online]. Available: <http://topology.neclab.eu/>
- [9] Internet topology collection, internet research lab der ucla. [Online]. Available: <http://irl.cs.ucla.edu/topology/>
- [10] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet AS-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, pp. 53–61, 2005.