

Measuring ISP Topologies with Rocketfuel

Dennis Knorr
(knorr@in.tum.de)

Betreuer: Stefan Kornexl

Hauptseminar „InternetMeasurement“ ,
Technische Universität München

Sommersemester 2005 (Version vom 7.Juni 2005)

Zusammenfassung

Dieser Artikel handelt von den Möglichkeiten, das Internet, oder seine Teilnetze, die großen Provider auf Routerebene zu kartieren. Hierzu wird die Software Rocketfuel genutzt. Diese erlaubt eine Reduzierung der benötigten Traces, ohne daß dabei die Genauigkeit der Karten nachläßt. Dazu werden BGP-Routing-Tabellen, Traceroute-Logs und DNS-Informationen benutzt. Schließlich werden die erzeugten Karten mit jenen, welche die Provider besitzen, verglichen, und damit die Effektivität der eingesetzten Verfahren analysiert.

1 Einleitung

Heutzutage werden genaue Karten des Internets immer wichtiger für Forschung und vielfältige Anwendungen. Die verschiedenen möglichen Wege durchs Internet, die Verbindungen nehmen können, beeinflussen Skalarität und Effizienz für alle Arten von Anwendungen und Protokollen. Auch Denial-of-Service-Angriffe können damit besser zurückverfolgt und analysiert werden. Leider existieren nur sehr wenig bis gar keine relevanten Karten, da dies zum Teil sicherheitssensitive Informationen sind. Da das Internet heutzutage sehr dynamisch und komplex ist, beschränken sich die Autoren von Rocketfuel darauf, die Struktur von Teilnetzen des Internets zu analysieren. Da sie eine Karte des Netzes auf IP-Ebene wollen, teilen sie die zu analysierenden Netze auf AS-Level auf. Dieser Artikel beschränkt sich darauf, die von Rocketfuel verwendeten Verfahren zu erklären, deren Effizienz mittels der Providerkarten zu bezeugen, und sie mit anderen Verfahren zu vergleichen.

Zunächst werden die Ideen von Rocketfuel[MITRWS] benannt. Ein Verfahren von Rocketfuel wird *Directed Probing* genannt. Es verwendet Routing-Informationen aus dem BGP-Protokoll, um Traceroute-Logs zu verwenden, die das Netz des von uns betrachteten Providers durchqueren. Die Technik des *Path Reduction* erkennt Routen von Verbindungen, welche einen ähnlichen oder gar den gleichen Weg durch ein Subnetz wählen. Allein diese beiden Verfahren reduzieren den Aufwand des Durchsuchens des Netzes im

Vergleich zur erschöpfenden Suche bei gleichbleibender Genauigkeit um ca. 3 Größenordnungen. Zu den bisherigen Techniken fügt Rocketfuel noch die *Alias Resolution* hinzu. Diese Technik zeigt auf, welche IP-Adressen zu einem Router gehören, wodurch es möglich wird, sehr viele IPs auf einige wenige physisch existierende Router abzubilden. Zum Schluß wird die Feinstruktur des Providers zusätzlich zu den bisher gewonnenen Erkenntnissen durch DNS-Informationen bestimmt. Diese bestimmt nämlich das Backbone und einzelne Subknoten des ISPs, sogenannte POPs (Point of Presence). Im nächsten Kapitel wird das Problem der Kartenerzeugung für Netze grundsätzlich erörtert. Im darauf folgenden Kapitel werden die Mapping-Techniken vorgestellt und näher erläutert. Im 3. Kapitel wird Auswirkung der oben genannten Verfahren genauer dargestellt. In Kapitel 7 wird gezeigt, wie die Rocketfuelsoftware aufgebaut ist, sowie im 8. Kapitel die Überdeckung der Karten von Rocketfuel und verschiedener kooperativer ISPs aufgezeigt werden. Danach werden noch Vergleiche zu anderen Kartierungsverfahren erläutert.

2 Probleme der Netzkartographie

Das Internet im gesamten ist ein sehr heterogenes System. Es gibt engvermaschte Bereiche, ebenso große Flächen der Erde, wo gar keine Netze existieren, und diese Gebiete sind dann auch nur über Satellit zu erreichen. Das zeigt auch schon die nächste Eigenschaft des Netzes auf. Die Verbindungsarten sind sehr unterschiedlich in Anforderung, Qualität, Anbindung und Erreichbarkeit. Zusätzlich ist es ein geschichtetes System, was die Transparenz, Dynamik und Heterogenität der Subsysteme stark fördert. Wenn man daher einen Graphen des Netzes erstellen will, muß man sich sehr stark an die Gegebenheiten anpassen, welche die Ausprägung des Netzes bestimmen. Hierbei kann man noch gesondert hervorheben, daß das Internet klassische Grenzen wie natürliche physische Grenzen oder staatliche Souveränität wenig beachtet. Vielmehr wird der Grad der Vernetzung durch wirtschaftliche Ballungsräume und Gebiete mit hoher Besiedelungsdichte bestimmt. Folglich kann man die Netze mittels der Provider einteilen, die ein oder mehrere Gebiete einzeln oder gemeinsam mit einer Anbindung an das Netz versorgen.

Beim Netz eines Providers unterscheiden wir dann sein Backbone sowie die unterschiedlichen POPs (Point of Presence). Ein POP ist eine lokale Gruppierung von Routern, die mit dem Backbone des ISPs verbunden sind. Man kann also das Backbone als die Verbindung aller POPs eines Providers definieren. Jede lokale Gruppierung von Routern des Netzwerks verwendet auch Systeme, die zur Routenorganisation BGP verwenden. Die meisten Router, die keine BGP-Systeme benutzen, gehören kleineren Organisationen wie z.B. kleinen Providern, welche die Anbindung zur letzten Meile bereitstellen.

Der erste Gedanke, um ein Netz vollständig zu erfassen, ist eine Verbindung von jedem Punkt zu jedem anderen Punkt des Netzes zu erstellen. Die theoretische Informatik zeigt hierbei jedoch, daß dieses Verfahren sehr ineffizient ist, da bei jedem Hinzufügen eines Knotens mit mehr als einer Verbindung zum Netz, der Algorithmus das Netz neu aufspannen müsste. Weiterhin wird hier der Tatsache nicht Rechnung getragen, daß viele Knoten im Netz, d.h. die Router, viel mehr Verbindungen haben, als die Endpunkte, welche die entfernten Systeme repräsentieren.

Um das Netz eines Providers zu unterteilen und kleinere Probleme zu betrachten, wird das Netz in seine AS-Bereiche aufgeteilt. Die AS-Nummern gewinnt man aus den BGP-Routing-Tabellen von Routeview[RTV]. Um in den Subnetzen genaueren Aufschluß über die Verbindungen der Router zu erhalten, werden Traceroute-Logs verwendet. Wenn man alle gewonnenen Traceroute-Daten zusammenwirft, erhält man schon eine, noch ungenaue, Karte des Netzwerks. Dazu werden bei Rocketfuel verschiedene Traceroute-Server verwendet, die mehrere Blickwinkel in die durch die AS-Bereiche definierten Subnetze bieten.

Doch auch der Ansatz, alle Präfixe der BGP-Tabelle mit den öffentlichen Traceroute-Servern anzuvisieren, würde die Server überlasten. Außerdem würde das Sammeln der Informationen zu lange dauern. Wenn man z.B. Traceroutes von allen Traceroute-Servern zu 120000 Netzpräfixen sammeln würde, und jeder Traceroute-Server eineinhalb Minuten braucht, um eine Anfrage zu beantworten, hätte man das Netz erst nach ca. 125 Tagen vollständig erfaßt (Das ist ein Drittel eines Jahres!). In dieser Zeit kann sich die Struktur jedes POPs und des Backbones von Providern stark verändern.

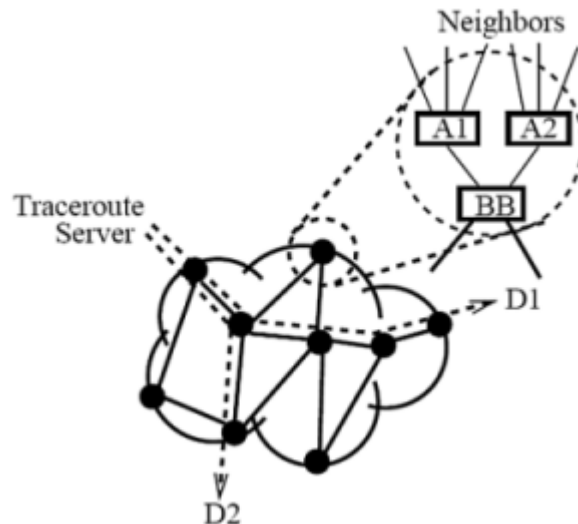


Abbildung 1: Beispielstruktur eines ISP mit POPs, Backbone und Nachbarn

Diese Abbildung ist, wie auch alle folgenden, [MITR] entnommen.

Der Gedanke bei Rocketfuel ist, daß Traceroute-Daten zwar die meisten Informationen über das Netz bereitstellen, allerdings sind sie auch diejenigen, welche am redundantesten sind. Als Konsequenz wird versucht, die Daten aus dem Bestand zu filtern, welche keine neuen Informationen über das Netzwerk liefern. Für eine Netzwerkkarte, die keine Redundanz enthält, muß man daher die Aliase der Router herausfinden. Da frühere Ansätze zur Aliasextraktion mangelhaft erschienen, wird die Routererkennung über eine Kombination von IP-Identifizier, Rate-limiting und TTL-Values-Verfahren bereitgestellt.

2.1 Traceroute - Was ist das?

Auch wenn Traceroute ein bekanntes Verfahren sein dürfte, wird es hier kurz dargelegt, wie es in [TCPILL] beschrieben ist. Für genauere Informationen kann man dort nachschlagen. Es gibt für das Verfahren der Erkundung eines existierendes Pfades mehrere Möglichkeiten. Da wäre IP Source Records, oder das Plazieren der IPs im Header. Am gebräuchlichsten erwies sich jedoch das Verfahren der TTLs in Verbindung mit ICMP. Traceroute schickt für ein bestimmtes Ziel eine UDP-Nachricht auf einem hohen Port (höher als 30000) mit einer TTL von 1 an das Ziel. Der nächste Router bekommt die Nachricht, dekrementiert die TTL und schickt an den Verursacher eine ICMP-Antwort "Time exceeded". Durch diese Antwort weiß der Host auch, was für eine Roundtrip-Zeit zu diesem Router existiert, und welchen Namen dieser hat. Danach wird das gleiche Ziel mit der TTL für 2 und 3 und alle TTLs getestet, solange bis der entfernte Rechner gefunden wurde. Danach hat man die Namen aller Router und die RTTs zwischen den beiden Rechnern. Wenn man von einem oder mehreren Rechnern alle IP-Adressen eines

Netzwerkes mit Traceroute getestet, hat man, bis auf eventuelle Backuprouten alle Verbindungswege des Netzes.

2.2 BGP - eine kurze Einführung

BGP, das Border Gateway Protocol ist dazu da, damit Router in autonomen Systemen Informationen über die Erreichbarkeit von Hosts inner- oder außerhalb ihres Netzes anderen Routern mitteilen können. Die Bereiche des Netzes werden in AS (Autonomous Systems) aufgeteilt, die über die ASNs, die Autonomous System Numbers, referenziert werden. Hierbei kann einem oder mehreren Netzwerkbereichen, auch Präfixe genannt, diese Nummer zugeteilt werden. Später können andere Router mit diesem Wissen Pakete in die entsprechenden Netzwerke korrekt weiterleiten.

3 Mappingtechniken

In diesem Kapitel werden Verfahren beschrieben, wie man Router entdecken kann, und wie diese für die verschiedenen AS-Bereiche eingeteilt werden. Weiterhin wird beschrieben, wie man Aliase der Router entdeckt.

3.1 Directed Probing

Directed Probing versucht die Traceroutes zu unterscheiden, die das Netzwerk des Providers durchqueren. Wenn man für jeden Eintrittspunkt ins Netzwerk die vollständige BGP-Tabelle hat, kann man den genauen Pfad durch das Netz für alle Verbindungen bestimmen. Da diese nicht vorhanden sind, werden die Daten von Routeview[RTV] als Annäherung verwendet.

Eine BGP-Tabelle bildet IP-Bereiche und deren Präfixe auf AS-Pfade ab, mit denen die IPs innerhalb der AS-Bereiche erreicht werden können. Dadurch ergeben sich 3 Klassen von Traceroutes, welche die AS-Pfade innerhalb des zu analysierenden Netzwerks betreffen.

- Tracerouten zu *abhängigen Präfixen* Die Präfixe bzw. Netzbereiche des ISPs oder eines seiner Subunternehmer nennt man abhängige Präfixe, wenn alle Tracerouten zu diesen IP-Bereichen von jedem Eintrittspunkt aus den ISP durchqueren. Das Wort Präfix rührt von der Einteilung der IP-Klassen bei Subnetzen her.
- Tracerouten von Insidern hat man, wenn der Traceroute-Server innerhalb eines IP-Bereiches des ISP steht. Die Traceroute von diesem Bereich zu allen anderen Präfixen durchquert den ISP.
- UpDownTraces nennt man Routen, die innerhalb des ISPs mittels der BGP-Routing-Tabelle bzw. dem AS-Pfad das Netz durchqueren. Sie bleiben innerhalb der schon bekannten Routen. Wenn man also den Weg zum Netzwerkbereich über die ASNs hat, und der Traceroute-Server in einem der autonomen Systeme auf dem Weg zum Netzwerkbereich liegt, ergibt das eine solche Route.

Einerseits beschleunigt *Directed Probing* das Mapping-Verfahren beträchtlich, andererseits entstehen dadurch relativ schnell auch Fehler. Zum einen werden eventuell Traceroute-Daten weggeworfen, die aber den Provider durchqueren, und zum anderen können in den Daten immer noch Tracerouten sein, die gar nicht unsere AS-Bereiche durchqueren. Erstere werden *false Positive*, zweitere *false Negatives* genannt.

3.2 Path Reduction

Nicht alle Traceroute-Verbindungen gehen durch unterschiedliche Wege im Netz des Providers. Mittels *Path Reduction* ist man in der Lage, die gleichartigen Verbindungsdaten weiter zu minimieren.

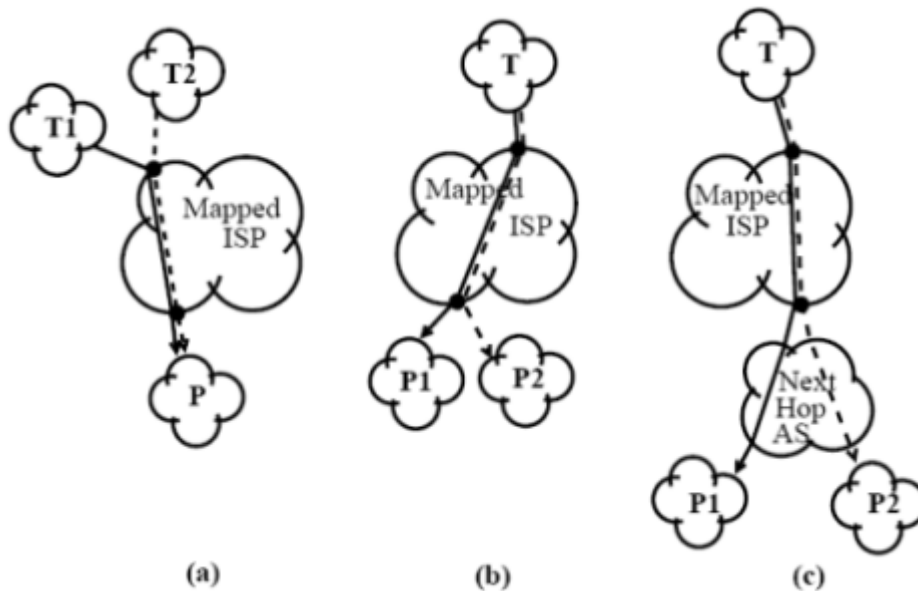


Abbildung 2: Wege durch das Netzwerk des Providers

- *Ingress Reduction*
Auch der Weg von IP-Paketen durch ein Netzwerk ist abhängig vom Ziel. Wenn daher 2 Verbindungen den gleichen Eintrittsknoten in das Netz und das gleiche Ziel haben, wenn sie das Netzwerk durchqueren, dann ist der Pfad innerhalb des Netzes des Providers der gleiche. In der Praxis kann man durch dieses Verhalten die Last der verwendeten Traceroute-Server minimieren.
- *Egress Reduction*
Ähnlich wie bei der *Ingress Reduction* haben Verbindungen, die das Netzwerk des ISP im gleichen Knoten verlassen, auch die gleiche Route innerhalb des Netzwerks genommen.
- *Next-hop AS Reduction*
Diese Technik ist ähnlich zur *Egress Reduction*, nur daß man hier zur Selektion der Traceroute-Daten nicht den Austrittsknoten wählt, sondern den nächsten AS-Hop, der zwischen dem ISP und dem Ziel liegt. Dieses Kriterium ist eine Verschärfung der Egress Reduction. In diesem Fall hier weiß man, daß es sogar nach dem zu analysierendem Netzwerk sehr wahrscheinlich noch den gleichen Weg noch etwas beibehält.

3.3 Alias Resolution

Trotz der Verringerung der vorhandenen Traceroute-Daten bleibt immer noch das Problem, daß man nicht weiß, welche IPs zum gleichen Router gehören. Bisher wurde dafür das Verfahren des Mercatorprojekts[HIMD] verwendet. Es verifiziert Alias-IPs durch Senden eines Traceroute-Probes auf einem hohen UDP-Port mit einer TTL von 255. Das

Verfahren gründet darauf, daß die Antwort des Routers durch die Standardkonfiguration *UDP Port unreachable* ist. Wichtig hierbei ist, daß die Antwort des Routers auch bei 2 Aliasen von der gleichen Quelladresse kommen. Da man zur Routererkennung nur 1 UDP-Paket braucht, ist dieses Verfahren zwar sehr effizient, allerdings wurden damit viele Aliase nicht gefunden. Das Verfahren bei Rocketfuel geht weiter, in dem es noch TTLs und ICMP-Rate-Limiting vergleicht.

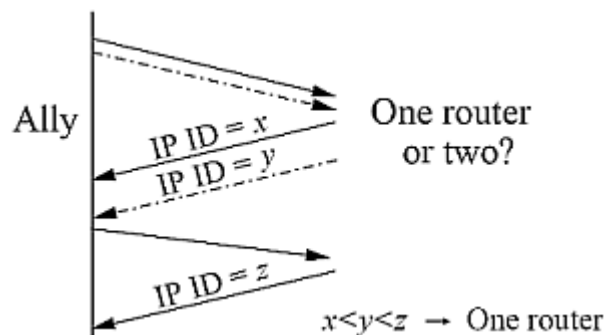


Abbildung 3: Wege durch das Netzwerk des Providers

Allerdings ist keine der vorher genannten Techniken so effektiv wie das Scannen über die sogenannten IP-Identifiers. Die Pakete eines Routers, die aufeinanderfolgend von ihm geschickt werden, haben auch die aufeinanderfolgenden IDs. Dabei geht man folgenderweise vor. Wenn man wie beim Mercatorprojekt[HIMD] 2 Antworten vom Typ Port unreachable bei unterschiedlichen IPs bekommen hat, sendet man an die erste IP noch ein drittes Paket. Sind die IP-Identifiers in der Reihenfolge grösser, gleich der Reihenfolge des Sendens, und die Differenz der IP-Identifiers ist sehr klein, kann man davon ausgehen, daß es der gleiche Router ist. Falls die IP-Identifier-Technik aufgrund des Rate-Limiting nicht funktioniert, erkennt die Heuristik des oben erwähnten ICMP-Rate-Limiting Tests die Aliase des Routers. Rate Limiting funktioniert auf relativ simple Weise. Die Testziele werden als erstes neu sortiert, wenn nur auf den ersten Ping eine Antwort kommt. Danach werden 2 Pings nach 5 Sekunden wieder gesendet. Wenn wieder bloß auf den ersten Ping geantwortet wird, allerdings diesmal von einer anderen Testadresse, erkennt die Heuristik einen Treffer.

3.4 DNS-Information über Router

Ein großer Vorteil von Filtern von Routern über die DNS-Einträge ist, daß man Kabelmodems, also Breitband-Dialup-Zugänge, recht leicht erkennt. Router, die nicht über BGP operierende Nachbarn haben, werden von den Providern auch häufig nach den AS-Bereichen selbst benannt. Durch diese Eigenart kann man dann auch die Grenzen des ISPs, auf den man sich konzentriert, sehr gut feststellen. Dies wird auch erleichtert, da die Provider oft Namenskonventionen haben, wodurch die Abgrenzung noch weit granularer gelingt.

4 Wirkung der Techniken

Die vorgestellten Techniken scheinen die Traceroute-Daten auf die wirklich wichtigen Daten zu minimieren. Um die Qualität dieses Verfahrens beurteilen zu können, sollen diese hier nochmal unter die Lupe genommen werden.

4.1 Directed Probing

Beim *Directed Probing* unterscheidet man 3 Fälle. Als erstes gibt es die Traceroute-Daten, die man wegwerfen kann. Dann gibt es die weggeworfenen Traceroute-Daten, die man zur Analyse hätte behalten sollen, und schließlich, als Gegenstück, die Daten, die in der analysierten Menge vorhanden sind, obwohl sie redundant zu anderen vorhandenen Traceroutes sind.

Die Effektivität von *Directed Probing* ist beachtlich. Ein Brute-force für die durch die BGP-Daten ermittelten Bereiche und den aufgeteilten ISP Präfixen der Größe /24 hätte minimal 90 Millionen Traceroute-Versuche gebraucht. Nun brauchen wir maximal 8% davon. Um einschätzen zu können, wieviele der nützlichen Traceroutes nicht verwendet wurden, wurde mit dem Projekt Skitter ein Experiment durchgeführt. Skitter ist ein anderer Ansatz zur Kartographierung von Netzwerken. Durch die Skitterdaten konnte berechnet werden, daß der Teil von Traceroutes, der nützlich gewesen wäre, aber weggeworfen wurde, zwischen 0,1 und 7% lag. Diese Spanne entsteht dadurch, daß innerhalb des ISP ein zufälliger Host als Ziel ausgewertet wird. Selbst bei 7% hätte man von 100 nützlichen Traceroute-Daten nur 7 Einträge verworfen.

Um Analysieren zu können, wieviele nutzlose Traceroute-Einträge wir behalten haben, müssen nur die Daten der Datenbank wiederholt mit den Reduktionsverfahren bearbeitet werden. Von allen Traceroute-Einträgen, die wir behalten haben, waren höchstens 6% nutzlos. Diese niedrigen Prozentzahlen zeigen, daß *Directed Probing* ein effektives Verfahren zur Reduzierung von Traceroute-Daten ist.

4.2 Ingress Reduction

Die Ingress Reduction sortierte 88% aller Traces aus, die *Directed Probing* behalten hatte. Indem bei der Erzeugung öffentliche Traceroute-Server verwendet wurden, gab es eine große Zahl von Startpunkten, in die von Rocketfuel gemappten AS-Bereiche. Durch die Tatsache, daß trotz vieler Knoten außerhalb des AS es nur eine kleine Zahl von Routern gibt, über die Pakete ins Netzwerk kommen, reduziert dies den Aufwand der nötigen Arbeit noch einmal.

4.3 Egress Reduction

Die Egress Reduction hat i.A. nur 18% der abhängigen Präfixe, die durch das *Directed Probing* erzeugt wurden, behalten. Es stellte sich auch heraus, daß man die Netzwerkbereiche in der Analyse in kleinere Bereiche aufspalten muß, da sonst eine beträchtliche Anzahl an Routern des ISP übersehen werden.

Um zu zeigen, daß diese Theorie, daß die Aufteilung auf 24-Bit Subnetze ausreicht, um Egress Router zu entdecken, werden zufällig 100 24er Bereiche aus den abhängigen Präfixen gewählt und diese in 30er Bereiche aufgeteilt. Nach dem diese Bereiche durchsucht wurden, merkte man, daß im Schnitt 8% mehr Egress Router gefunden wurden. Aus diesem Grunde wußte man, daß Egress Reduction nur für einige ISPs gute Ergebnisse brachte, was durch die Anzahl der Subprovider bedingt sein kann. Deshalb werden die Bereiche in Zukunft dynamisch allokiert und getestet werden, ob diese Netze in kleinere Teile aufgeteilt werden sollen.

4.4 Next-Hop AS Reduction

Diese Reduktion verringert das Datenaufkommen von UP/Downtraces auf 5% dessen, was *Directed Probing* erbrachte. Next-Hop Reduction ist so effektiv, weil die Anzahl der AS-Bereiche im nächsten Hop viel kleiner ist, als die Anzahl der Präfixe. Dies ist für

Insider sehr wichtig, die sonst 120000 Präfixe in der BGP-Tabelle eruieren müßten. In der Praxis stellte sich heraus, daß diese Reduktion in 7% der Fälle falsch lag.

4.5 Bewertung der Reduktionen

Dadurch, daß bei jeder Reduktion völlig andere Schwerpunkte gesetzt werden, ist die die Vereinigung dieser Techniken sehr effektiv. Man muß nur 0,1 % der Traces in Betracht ziehen, entgegen dem Brute-force-Ansatz. Die Streuung lag hierbei dann von ISP zu ISP variierend von 0,03 bis 0,05 % der Versuche.

Diese Kartierungsverfahren skalieren mit der Anzahl der Blickwinkel der Traceroute-Server, die wir auf den ISP werfen. Je mehr davon gegeben sind, desto schneller erhält man bessere Karten des Netzes, ohne die Anzahl der Traces erhöhen zu müssen. Neue Startknoten bzw. Blickwinkel verschnellern die Methode, oder machen die Karte genauer. Das hängt davon ab, ob der neue Startknoten sich einen Router mit einem anderen Startknoten teilt, oder ob dieser Router ein neuer Einstieg in das Netz des ISPs ist.

4.6 Alias Resolution

Das IP-identifizier-Verfahren mit Ally identifizierte fast 3 mal so viele Routeralias, wie das Verfahren von Mercator. Weiterhin waren die gefundenen Aliase von Mercator eine abgeschlossene Untermenge der Menge von Ally, wodurch man weiß, daß es ausreicht, wenn man bloß das Verfahren von Ally benutzt. Zum Testen verwendeten wir die von Ally gefundenen Aliase, durch die man durch die DNS-Namen erraten kann. Bei den 2 Tests mit den ISPs Ebone und Sprint ergab sich, daß bei Sprint 240 von 300 Routern erkannt wurden, und bei Ebone 119 von 139 Routern erkannt wurden. Allerdings tauchte das Problem auf, daß ein signifikanter Teil von IPs nie auf unsere *Alias Resolution* Anfragen antwortete. Die maximale Anzahl von Aliasen, die beobachtet wurde, waren 24, von einem AT&T-Router in New York.

5 Rocketfuel - Die Software

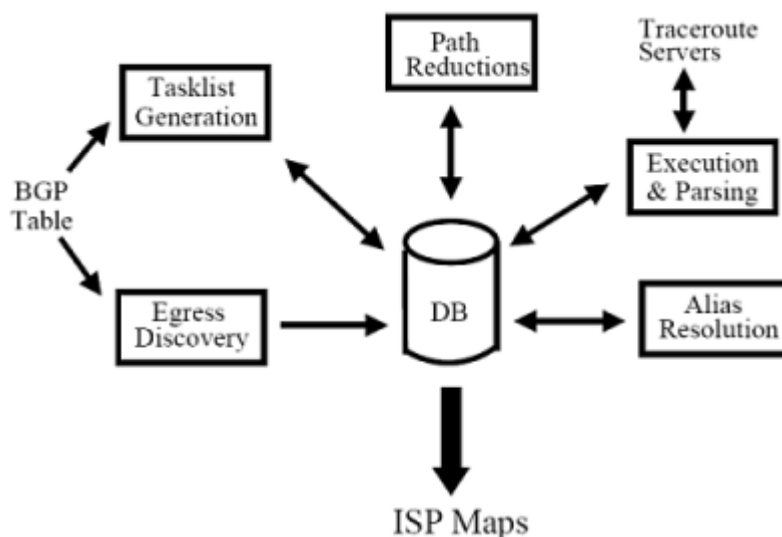


Abbildung 4: Struktur der Rocketfuelsoftware

Rocketfuel besteht in seiner Struktur aus einer Datenbank und mehreren Analysemodulen. Die Datenbank ist ein PostgreSQLserver, die für die konsistente Datenspeicherung sorgt und die Interprozesskommunikation der asynchron laufenden Module abgleicht. Es wurden öffentlich verfügbare Traceroute-Server, die bei traceroute.org gelistet sind, verwendet, um die Traceroute-Daten zu erhalten. Davon gibt es auf der ganzen Welt circa 784 Startpunkte auf die zu erforschenden Netzwerke. Es ist hier auch nicht unwahrscheinlich, daß ein Traceroute-Server viele Punkte in einem AS erreicht. Die BGP-Routing-Tabellen wurden von dem Projekt Routeview[HIMD] übernommen.

Egress Discovery ist das Modul, das Egress Router für abhängige Präfixe sucht. Um diese Ausgangsrouten zu finden, wird von der lokalen Maschine zu jedem abhängigen Präfix ein Traceroute durchgeführt. Danach werden die Präfixe in 24er Bereiche aufgeteilt. Das *Tasklist generation module* erzeugt durch die BGP-Tabellen eine Liste für die *Directed Probes*. Die abhängigen Präfixbereiche werden durch die Ausgangsknoten ersetzt, und Duplikate werden dabei gelöscht. Dadurch findet man genau den Pfad bis zu den Ausgangsknoten. Das Modul *Path Reduction* nimmt die Tasklist aus der Datenbank, wendet die Next-hop-AS und Ingress-Reduktion an, und erzeugt dann Jobs bzw. Traceroute-Befehle, die ausgeführt werden müssen. Wenn ein Traceroute ausgeführt wird, überprüft das Modul, ob der Einstiegs- oder Ausgangs-Router schon vorkam. Wenn ja, wird der Job aus der ausführungsliste genommen.

Das Execution-Modul verwaltet die Anfragen an die Traceroute-Server und arbeitet mit Loadlimiting und Loadbalancing. Load-distribution wird durch die Randomisierung der Jobliste erreicht, so daß jeder Traceroute-Server nur einmal innerhalb von 5 Minuten angesprochen werden muß. Der Traceroute-Parser holt die IP-Adressen, welche die Router-Interfaces und die Zwischenstationen darstellen, aus dem Output des Traceroute-Servers.

Das *Alias Resolution* Modul mit dem Ally-Verfahren braucht ein paar Tricks. Der Suchraum wird durch 3 Regeln begrenzt. Als erstes nutzt man die Hierarchie der DNS-Namen, indem man die IP-Router-Adressen und ihre aufgelösten Namen sortiert. Weiterhin könnten Router-IPs, die ähnliche TTLs zurückgeben, auch Aliase desselben physischen Routers sein. Es werden immer die IPs getestet, deren TTL-Abstand am kleinsten ist. Erst die mit TTL-Differenz gleich Null, dann mit Eins, und so weiter. Schließlich wird natürlich ausgenutzt, daß Aliase transitiv sind. Dadurch bekommen wir immer mehr Aliase, wenn diese sich auflösen. Das Modul ist fertig, wenn bei allen Paaren von IP-Adressen die Aliase aufgelöst sind, oder nicht, oder die IPs nicht antworten.

6 Testergebnisse

In der Struktur der Backbones stellten sich zwischen den ISPs und den kleineren natürlich Unterschiede heraus. Zum Beispiel haben zwar AT&T und Sprint beide größere POPs in den größeren Städten, allerdings ist Sprint im Hinterland lange nicht so stark vertreten wie AT&T. Auch verwenden jüngere Provider im Backbone nicht mehr nur die traditionelle Technologie, sondern auch immer mehr MPLS, ATM oder frame relay PVCs um die einzelnen POPs zu verbinden. Die einzelnen POPs der Provider hingegen sind sich relativ ähnlich. Die lokalen Backbone-Router sind mit denen anderer POPs verbunden, und die Access-Router des POPs sind wiederum an die Backbone-Router angeschlossen.

Einige der Access-Router sind mit anderen lokalen Routern verbunden, manchmal jedoch werden diese einfach über Layer2-Geräte mit den benachbarten Bereichen zusammengeschlossen. Es ist, wie auch schon in den vorigen Kapitel erwähnt, durchaus üblich, daß die Backbone- und Neighbor-Router unterschiedlich der Firmenkonvention entsprechend gelabelt sind. Dadurch wird auch der Wert dieser DNS-Namen zur Topologieerstellung ersichtlich.

AS	Name	ISP		with customer & peer		POPs
		Routers	Links	Routers	Links	
1221	Telstra (Australia)	355	700	2,796	3,000	61
1239	Sprintlink (US)	547	1,600	8,355	9,500	43
1755	Ebone (Europe)	163	300	596	500	25
2914	Verio (US)	1,018	2,300	7,336	6,800	121
3257	Tiscali (Europe)	276	400	865	700	50
3356	Level3 (US)	624	5,300	3,446	6,700	52
3967	Exodus (US)	338	800	900	1,100	23
4755	VSNL (India)	11	12	121	69	10
6461	Abovenet (US)	367	1,000	2,259	1,400	21
7018	AT&T (US)	733	2,300	10,214	12,500	108

Abbildung 5: Tabelle für Leistungsfähigkeit/Anzahl der gefundenen Knoten

6.1 Abgleich der Ergebnisse mit den ISPs

Die Mapping-Techniken wurden bei vielen unterschiedlichen Internet Providern getestet. Darunter waren AT&T, Ebone, Exodus, Level3, Sprint, Tiscali und VSNL. 3 der 10 Provider die wir untersuchten, halfen, die Testergebnisse auf ihre Richtigkeit zu überprüfen. Es wurden auch die Teile des Netzes ermittelt, die bei der Kartierung mit Rocketfuel unentdeckt blieben. Bei dem Abgleich wurden folgende Ergebnisse bekannt:

- Kein POP der 3 Provider wurde von Rocketfuel übersehen.
- Die Verbindungen der POPs untereinander wurden alle erkannt.
- Bei zufällig gewählten POPS wurden zum Teil Access-Router nicht gefunden, oder fälschlicherweise Router vom benachbarten AS miteinbezogen.
- Die Frage, wie groß der Anteil von Routern von Subunternehmern war, die von Rocketfuel nicht erkannt wurden, konnten oder wollten die Vertreter der ISPs nicht beantworten.
- Die Vertreter schätzten die erzeugten Netzwerkkarten von gut bis exzellent ein.

Um die Kartierung zu vervollständigen, wurde in den Netzwerkpräfixen nach weiteren antwortenden IP-Adressen gesucht. Wenn bei dieser exhaustiven Suche weitere Router gefunden worden wären, wäre klar geworden, daß die benutzten Traceroute-Daten nicht alle Teile des Netzes abdecken. Dazu wurden zufällig 60 24er Bereiche jedes ISPs ausgewählt, die mindestens 2 Router beinhalten, um nach neuen Routern suchen zu können. Schließlich ist klar geworden, wenn der Provider sein Netzwerk logisch aufbricht und benennt, man zwischen 64%-96% der Backbone-Router finden kann. Die Abdeckung der Access-Router ist zwar nicht so gut, wie bei Backbones, aber immer noch beachtlich.

6.2 Vergleiche mit anderen Projekten

Obwohl Routeview mit BGP-Informationen arbeitet, ist es interessant zu sehen, daß Rocketfuel einige Neighbor-Router findet, die Routeview nicht erkennt. Allgemein gesehen kann man sagen, daß Rocketfuel eher größere Nachbarn sieht, und Routeview mehr kleinere Nachbarn erkennt. Mit "klein" ist hier gemeint, daß der AS-Graph des Netzwerkes einen niedrigen Grad hat.

Skitter ist ebenfalls ein Traceroute-Daten-basierendes Projekt, das von CAIDA betrieben wird. Die Tabelle zeigt, daß Rocketfuel sieben mal so viel Router, IPs und Verbindungen wie Skitter findet. Dafür erkennt Skitter aber Router, die Rocketfuel nie sieht.

7 Zusammenfassung

In dieser Arbeit wurden einige neue Werkzeuge vorgestellt, mit denen man die Struktur von großen Netzwerken erforschen kann. Sie haben den Aufwand im Vergleich zu Brute-force-Ansätzen um 3 Größenordnungen verkleinert, wobei nur wenig von der Genauigkeit der Karten verloren ging. Dies war nur mit Hilfe der öffentlichen Traceroute-Server möglich. Die neuen Werkzeuge wie die Alias-Resolution und die Pfad-Reduktion ermöglichten es, 10 ISPs zu kartieren. Um die Richtigkeit dieser Karten zu kennen, wurden sie mit den Karten verglichen, welche die Netzwerkadministratoren besaßen. Weiterhin wurden die Router, die Subnetze und die Peerings der einzelnen Provider gesucht und analysiert. Rocketfuel bewährte sich hierbei sehr gut. Trotzdem kratzt dieser Artikel nur an der Oberfläche, und die vorgeschlagenen Ansätze können sehr wahrscheinlich erweitert werden.

Literatur

- [MITR] Neil Spring, Ratul Mahajan, David Wetherall: *Measuring ISP Topologies with Rocketfuel*; University of Washington.
- [HIMD] Ramesh Govindan, Hongsuda Tangmunrakit: *Heuristics for Internet Map Discovery*; USC Information Sciences Institute
- [TCPILL] Richard W. Stevens: *TCP/IP Illustrated, Volume 1 The Protocols*; October 1993
- [MITRWS] <http://www.cs.washington.edu/research/networking/rocketfuel/>
- [RTV] <http://www.routeviews.org>