

Locality of Internet Connections

Jan-Pascal van Best
janpascal@vanbest.org

Willem G. Vree
w.g.vree@tbm.tudelft.nl

Delft University of Technology, Jaffalaan 5, 2628 BX Delft
The Netherlands

April 24, 2004

Abstract

Internet Service Providers (ISPs) interconnect with each other in order to deliver Internet services to their customers. By using data from RIPE's Test Traffic Measurements project, this project produced a map of the Internet's topology, which was subsequently transformed to a map showing the interconnections between ISPs. This map is visualised using a simulated annealing approach (minimising the number of link crossings in the map). For this visualisation highly connected ISPs, usually global ISPs like UUNet or Cable and Wireless, are left out of the map. The resulting map shows a high degree of locality of interconnections, indicating that small and medium sized ISPs favour interconnecting with local partners.

1. Introduction

The research reported in this paper is part of a project called UMEEPI: Understanding and Modeling of End-to-End Performance of the Internet. The project is a collaboration of the University of Amsterdam, Delft University of Technology, RIPE NCC and KPN Research. The focus of the project is on the analysis of packet dynamics derived from the specific active measurements performed by the RIPE Test Traffic Measurements (TTM) project.

The first step we undertook was to visualise the large collection of data that has been assembled by RIPE NCC in the past three years. From these visualisations we hope to get clues for further modeling and analysis. This paper reports a remarkable fact that became apparent as soon as we could visualise the routes that IP packets follow on the level of autonomous systems (ASs). We discovered that it was possible to obtain close to planar embeddings of the AS graphs when only a small part of the ASs was left out. To make these pictures, we devised a specialised version of a simulated annealing algorithm. A second discovery was that the (geometrical) locality in these graphs corresponded to real geographical proximity. Combined with the interpretation that AS graphs correspond to peering relations between ISPs we came to the conclusion that smaller ISPs favor peering with local partners.

To support these claims we present three graphs of the Internet as perceived by the RIPE NCC test boxes with a period of one year in between. For each of these graphs the annealing results are consistent and show the geographical locality of around 85% of the observed ISPs.

In chapter 2 we present some background of the RIPE Test Traffic project and the quality of the data. In chapter 3 we explain how IP numbers are mapped to autonomous systems. Chapter 4 outlines the details of our graph drawing algorithm by simulated annealing, and in chapter 5 the resulting graphs and conclusions are presented.

2. Topology Data

The results presented in this paper are based on data collected by the RIPE Test Traffic Measurements project (TTM). The goal of TTM is to perform independent measurements of connectivity parameters, such as delays and routing vectors, from the Internet[UKKW98]. The metrics used are the ones being developed by the IPPM (Internet Protocol

© Copyright 2002 Jan-Pascal van Best and Willem G. Vree

This research was made possible in part by a grant from the Dutch Ministry of Transport, Public Works and Water Management

Performance Metrics) working group and consist of end-to-end unidirectional delay, loss, and routing of UDP data packets[AKZ99, IPPM, PAMM98].

To perform these active one-way measurements RIPE NCC has developed dedicated test boxes. Although at present approximately 60 test boxes have been installed at ISPs [RTT], the number of active test boxes observed in the three year period of our analysis lies between 40 and 50. Most ISPs are located in the geographical area served by RIPE and most test boxes are located close to a border router of the ISPs network. This means that the delays and routing information that is observed relates mainly to the backbone of the European part of the Internet. The TTM Test Box has been designed with a lot of attention paid to security, as the data is gathered at critical points in the backbone networks of ISPs.

The analysis presented in this paper only uses the routing vectors of the TTM data (not the delay and loss data). The routing vectors are collected by all test boxes with an improved version of the traceroute tool[UKKW98, KZ99]. In the data sample of the first 12 days of March 2002 we found 23590 routing vectors containing 364063 IP addresses (hops). Only 2583 of these IP addresses did not respond to the traceroute tool. Thus only 0.7% of the hops have an unknown IP address, which is a quite good figure compared to what can be expected from UDP-probing. The unknown addresses have been left out of the maps presented in the following chapters. The fact that we present AS level graphs further alleviates the problem of unknown addresses.

Each of the three graphs that we show in this paper is based on twelve snapshots of the TTM test box data, taken from the 1st to 12th of March in, respectively, 2000, 2001 and 2002. We have taken these twelve snapshots with an interval of one day to be sure to capture most ASs without running into troubles caused by changing routes and possibly changing peering arrangements. On the scale of twelve days the AS-graphs appear to be largely static and, on the other hand, few ASs that normally should route test box packets will remain unobserved (which occasionally happens due to technical problems with routers or test boxes) for twelve consecutive days.

Each of the twelve snapshots is taken at the same (busy) time of the day (15:00). With these precautions we assume that the graphs represent a fairly complete AS-picture of the Internet as observed by the TTM test boxes in the beginning of March during three consecutive years.

Two other Internet mapping projects should be mentioned here. At Advanced Network & Services in the USA a similar effort to the RIPE TTM project is undertaken, called Surveyor[KZ99]. Though both projects differ in technical implementation, their measurements are the same. The CAIDA organisation tries to track and visualise the Internet, using the SKITTER-tool[HPDC02]. CAIDA aims for complete coverage of the Internet and takes an asymmetric approach using active probing with much more targets than monitors and using two-way measurements.

3. Mapping to Autonomous Systems

An Autonomous System (AS) is defined as a connected group of one or more IP prefixes run by one or more network operators which has a single and clearly defined routing policy[HB96]. An AS is usually controlled by an ISP, a university, a business enterprise, or a business division. An autonomous system is assigned a globally unique number, an Autonomous System Number or ASN. Since ASs can be seen as units of Internet routing, and because an AS is not just a networking entity, but also a real organisation, mapping the Internet on the level of Autonomous Systems gives us insight into the *business* network of Internet service provision.

Chang et.al. [CJW01] describe a method for mapping IP addresses to AS numbers. This method makes use of two sources of information: BGP routing tables and the Internet Routing Registry[IRR]. The information from the IRR is more commonly known as “whois data”, because the IRR can be queried using the `whois` tool. In our AS mapping process we have used only whois data, since most of the IP addresses in our data sets are from the RIPE area, and the RIPE IRR database is fairly up-to-date[CJW01]. We have made use of a list of 20 whois servers, including `whois.ripe.net`, `rr.arin.net`, `whois.apnic.net` and `whois.radb.net`. Less than 1% of the IP addresses in our data set could not be mapped onto an AS number with this method. These unmappable IP addresses have been left out of further analysis.

All IP addresses in the routing vectors taken from the TTM data are mapped onto autonomous systems. Directed edges are added from one AS to another if they appear consecutively in a (translated) routing vector. Edges in the resulting AS graph thus represent traffic flowing from an AS to another, indicating peering or transit relationships between those ASs.

4. Visualisation using Simulated Annealing

The graphs that result after mapping to autonomous systems still count around 120 nodes and 650 edges. Our first goal was to draw this graph in an aesthetically pleasing way. To this end, we have used the *simulated annealing* method [DH96]. This graph drawing method works by defining an energy function, dependent on the placement of nodes and edges in the graph. During the annealing process, individual nodes are moved tentatively. If that move decreases the total energy of the graph, the move is carried out; if the move increases the total energy, it is carried out with a probability that depends on the energy difference because of the move and on a *temperature*, that decreases in time. This whole procedure may be compared to “shaking” the graph, starting vigorously, and in time more and more gently.

1. Move a part of the nodes, with the highest degrees, to the right of the graph and do not take these into consideration in the rest of the algorithm. This makes drawing the rest of the graph much easier. The part retained in the annealing algorithm is named α .
2. Place the remaining, low-degree nodes randomly.
3. Initialize the temperature T to T_{start} .
4. Start with the first node of the graph and call it n .
5. Choose a random new location for node n . In order to speed up the algorithm, the displacement in the x and y directions are chosen randomly from the interval $\left[-\sqrt{\frac{T}{2\lambda_3}}, \sqrt{\frac{T}{2\lambda_3}} \right]$.
6. Calculate E , the energy difference because of the moved node. A positive E means an increase in the total energy because of the move.
7. If E is negative, move the node. If E is positive, move the node with a probability $p = e^{-\Delta E/T}$.
8. Repeat steps 5-7 for all nodes n .
9. Repeat steps 4-8 30 times in order to get to an equilibrium situation for this temperature.
10. Decrease temperature T with a fixed factor γ .
11. Repeat steps 4-10 until the final temperature T_{end} is reached.

Table 1: Annealing Algorithm

The algorithm is presented in more detail in Table 1. The difference between our method and the one described in [DH96] lies in step 5 of the algorithm. In the original method, a node is placed on the perimeter of a circle around its original position, with the circle's radius decreasing in time. In our algorithm, we have simplified this step in that we place the node inside a square with the size of the square decreasing in time. In step 5, the size is chosen in such a way that, at the current temperature, displacements outside of the square would not be allowed because of the “Edge Stretch” contribution to the energy function.

In step 6 of the algorithm in Table 1, the energy difference is calculated between two graph layouts, which differ in the location of a single node. The energy function used to calculate the difference between the two layouts is the following:

$$E = \sum_{\text{node } i, j} E_{ij}^{\text{NR}} + \sum_{\text{node } i} E_i^{\text{BR}} + \sum_{\text{edge } k} E_k^{\text{ES}} + \begin{cases} \sum_{\text{node } i, \text{edge } k} E_{ik}^{\text{NER}}, & \text{if } T \leq T_C \\ \sum_{\text{edge } k, l} E_{kl}^{\text{EC}}, & \text{if } T > T_C \end{cases}$$

This energy function used in our annealings consist of the components listed in Table 2. Below a certain temperature T_C , the algorithms starts *finetuning*, using the computationally intensive E^{NER} in stead of the simpler E^{EC} . λ_1 to λ_5 are parameters that determine the relative strengths of these components.

Node Repulsion	$E_{ij}^{NR} = \frac{\lambda_1}{d_{ij}^2}$	For every pair i,j of nodes, with d_{ij} the (euclidian) distance between the nodes. This contributes a repulsive force between nodes.
Border Repulsion	$E_i^{BR} = \lambda_2 \left(\frac{1}{x_i^2} + \frac{1}{(w-x_i)^2} + \frac{1}{y_i^2} + \frac{1}{(h-y_i)^2} \right)$	For every node i , with x_i, y_i the co-ordinates of the node, and w, h the width and height of the bounding box. This contributes a repulsive force between each node and each border.
Edge Stretch	$E_k^{ES} = \lambda_3 d_k^2$	For every edge k , with d_k the length of the edge. This contributes an attractive force between the endpoints of every edge
Edge Crossings	$E_{kl}^{EC} = \begin{cases} \lambda_4, & \text{if edges } k, l \text{ cross} \\ 0, & \text{otherwise} \end{cases}$	For every pair k, l of edges. This penalizes edge crossings more crudely than node/edge repulsion, but is more easily calculable, so this is used in the initial part of the algorithm.
Node/Edge Repulsion	$E_{ik}^{NER} = \max\left(\lambda_4, \frac{\lambda_5}{d_{ik}^2}\right)$	For every node i and edge k , with d_{ik} the distance between node i and edge k . This contributes a repulsive force between each node and each edge, which also hinders edge crossings. This energy component is used in the finetuning part of the algorithm.

Table 2: Energy function components

The parameters of the algorithm and of the energy function are listed in Table 3 with the values used in our annealings. The value of λ_3 is much smaller than that of the rest of the parameters because λ_3 is multiplied with a distance-squared, while most of the other parameters are *divided* by a distance-squared. When in finetuning mode, distances between node and edges smaller than the square root of λ_5/λ_4 are regarded as (edge) crossings. This is implemented in the formula for E^{NER} .

Parameter	Value used	Description
T_{start}	20	Initial temperature of the algorithm
T_{end}	0.00000005	Ending temperature of the algorithm
γ	0.85	Cooling factor
T_C	0.00005	Below T_C , the algorithm finetunes, using E^{NER} in stead of E^{EC}
α	0.8	Part of the (lowest-degree) nodes retained in the algorithm
λ_1	1	Strength factor of the node repulsion component of the energy function
λ_2	1	Strength factor of the border repulsion component of the energy function
λ_3	0.00000001	Strength factor of the edge stretch component of the energy function
λ_4	1	Strength factor of the edge crossings component of the energy function
λ_5	0.1	Strength factor of the node/edge-repulsion component of the energy function

Table 3: Algorithm Parameters

The results of annealing a typical AS-level graph using these settings can be found in Figure 1. The emphasis on preventing edge crossings in the algorithm takes care of producing aesthetically pleasing graphs. Layouting graphs of this type by hand typically results, after much labour, in one or two more edge crossings than what the algorithm provides.

The Java tool that has been developed for this research can be downloaded from [TRACES]. The tool can extract data from RIPE NCC's TTM database (or a mirror), map the resulting routing vectors to ASs using whois server data and perform the annealing layout algorithm detailed this chapter. Access to the TTM database is necessary to use this tool in a useful way.

5. Country Mapping and Results

The CAIDA organisation offers a service to determine the geographical location of IP- and AS numbers[MPDC00]. We have used this excellent service to map AS numbers to geographical positions.

Where geographical location of IP numbers is often rather difficult to derive, the location of most AS numbers is quite reliably recorded in the databases of RIPE, ARIN and APNIC. However, what is the meaning of the geographical location of an AS number?

The location of an AS mostly refers to the address of the head office of the owner of the AS, generally a particular ISP. In the graphs presented in this paper we only use the country part of this AS address. Our aim is to show the relation between ASs observed by the RIPE TTM test boxes and the countries where the head offices of the owners of these ASs are located.

Edges between ASs correspond to peering agreements and these agreements may be determined by geographical, cultural and legal influences. Knowledge of the countries where these arrangements are made is relevant, because most countries do have specific legal or cultural properties. What we hope to discover is whether AS graphs show relations during several years that suggest a geographical or cultural influence.

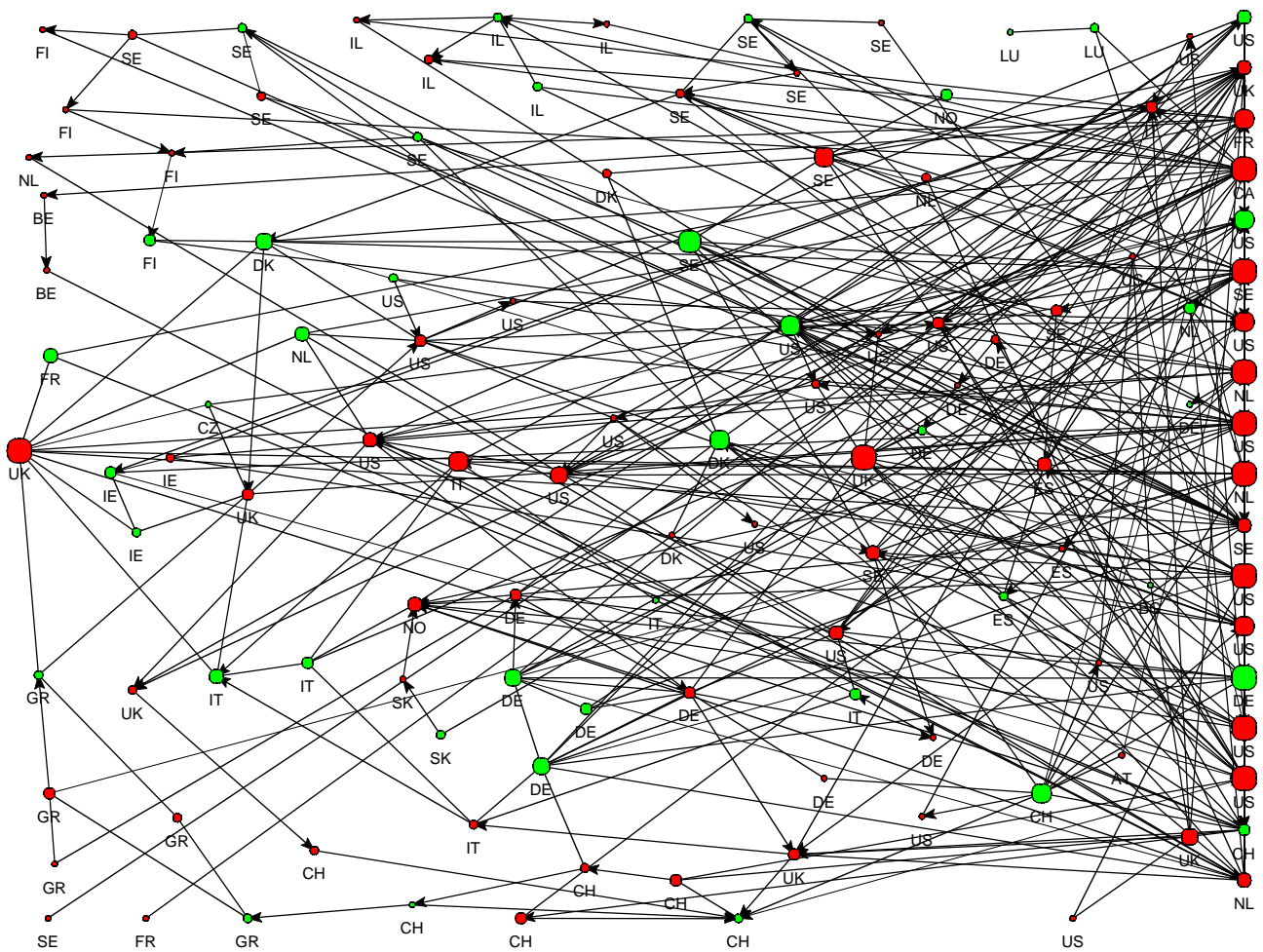


Figure 1: All ASs in March 2000

When plotting AS graphs we discovered that relatively few ASs (less than 16%, see Table 4) have a very large peering degree. For our purpose there are two reasons for not considering and drawing these large ASs. On the one hand they completely spoil a clear, planar layout of a graph, on the other hand one cannot expect from these ASs to see a lot of evidence for our search for geographical or cultural relations. Large ASs have peering relations with so many other countries that one is not likely to see any particular geographical or cultural influence.

Leaving out these few large ASs appears to result in graphs, still containing typically 84% of all observed ASs, which can be drawn with relatively few crossings using the simulated annealing algorithm discussed in the previous chapter. From this clear geometrical picture of the graphs we may be able to see geographical or cultural influence during the years of our observation.

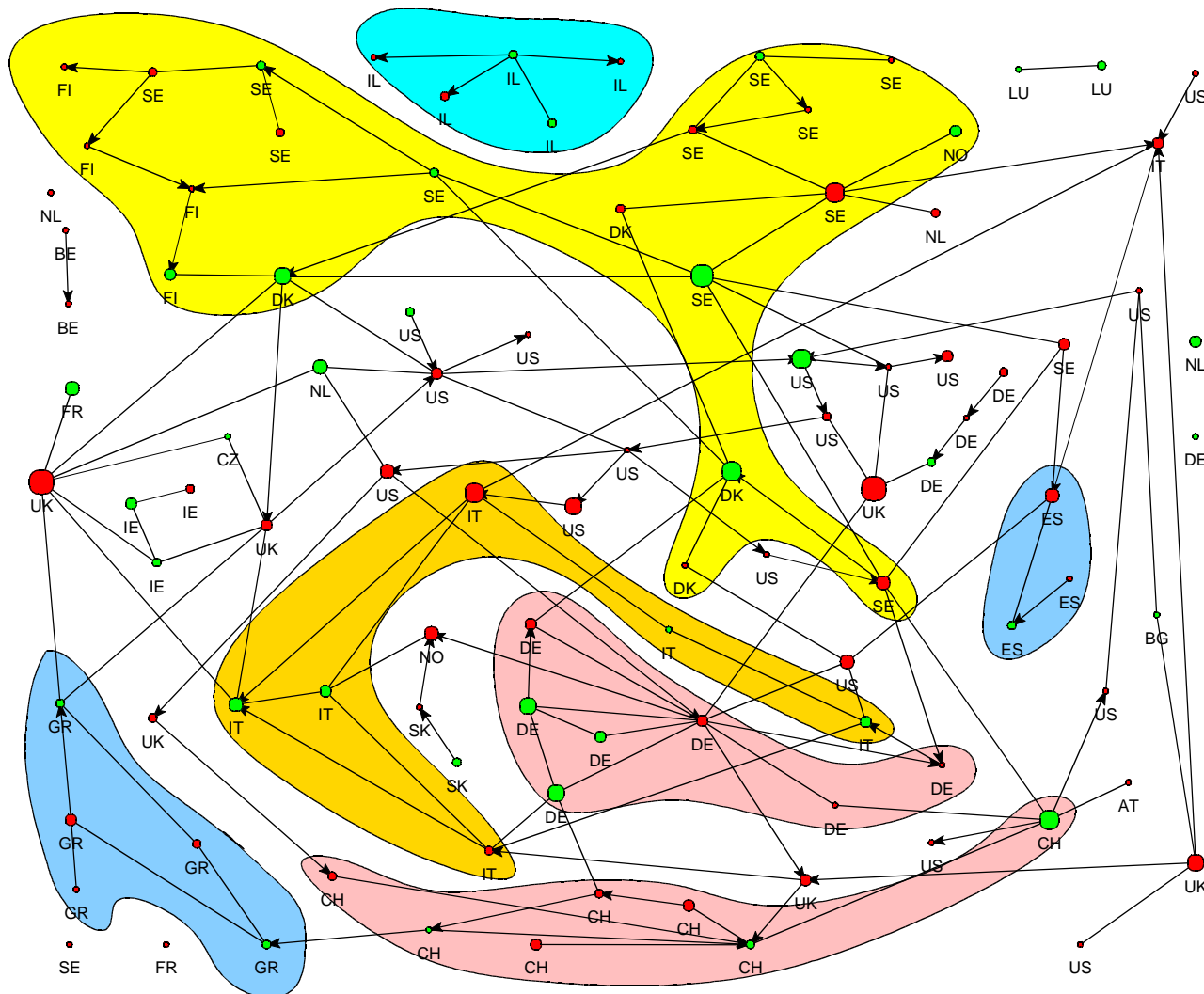


Figure 2: Small ASs in March 2000

Figure 1 shows the graph of all ASs observed in twelve snapshots taken from the 1st to the 12th of March 2000. The 16% largest ASs, having a peering degree of more than 16, are shown at the right border, positioned on a straight vertical line. The rest of the ASs is positioned with the simulated annealing algorithm. Figure 2 shows the same graph where the 16% largest ASs have been omitted, thus revealing a very clear picture of the 84% remaining smaller ASs in March 2000. In precisely the same way the graphs of the Internet in March 2001 and 2002 are shown in Figure 3 and Figure 4. Numerical data of the graphs is summarised in Table 4.

In the graphs edges without an arrow are bi-directional connections. The presence of an arrow means that packets only followed that single direction. Each edge may represent a single connection between two IP interfaces, or many connections between many interfaces, but all located in the two ASs connected by the edge.

The size of the nodes is proportional to the number of IP interfaces observed in the respective ASs. To be more precise: the radius of a node is proportional to the cube root of the number of IP-interfaces. So large nodes in the graph are a lot larger than small nodes. Looking at the graph immediately conveys a feeling for the relative size of the ISP's.

A lightly shaded node (green) contains one or more test boxes, whereas a darkly shaded node (red) contains no test box. In this way the pictures give a quick impression of how test boxes are distributed and how many there are in relation to the other (dark) nodes.

The graphs show some isolated nodes or isolated small sub-graphs. This just means that the peering partners of these nodes or sub-graphs are in the omitted set of 16% large ASs.

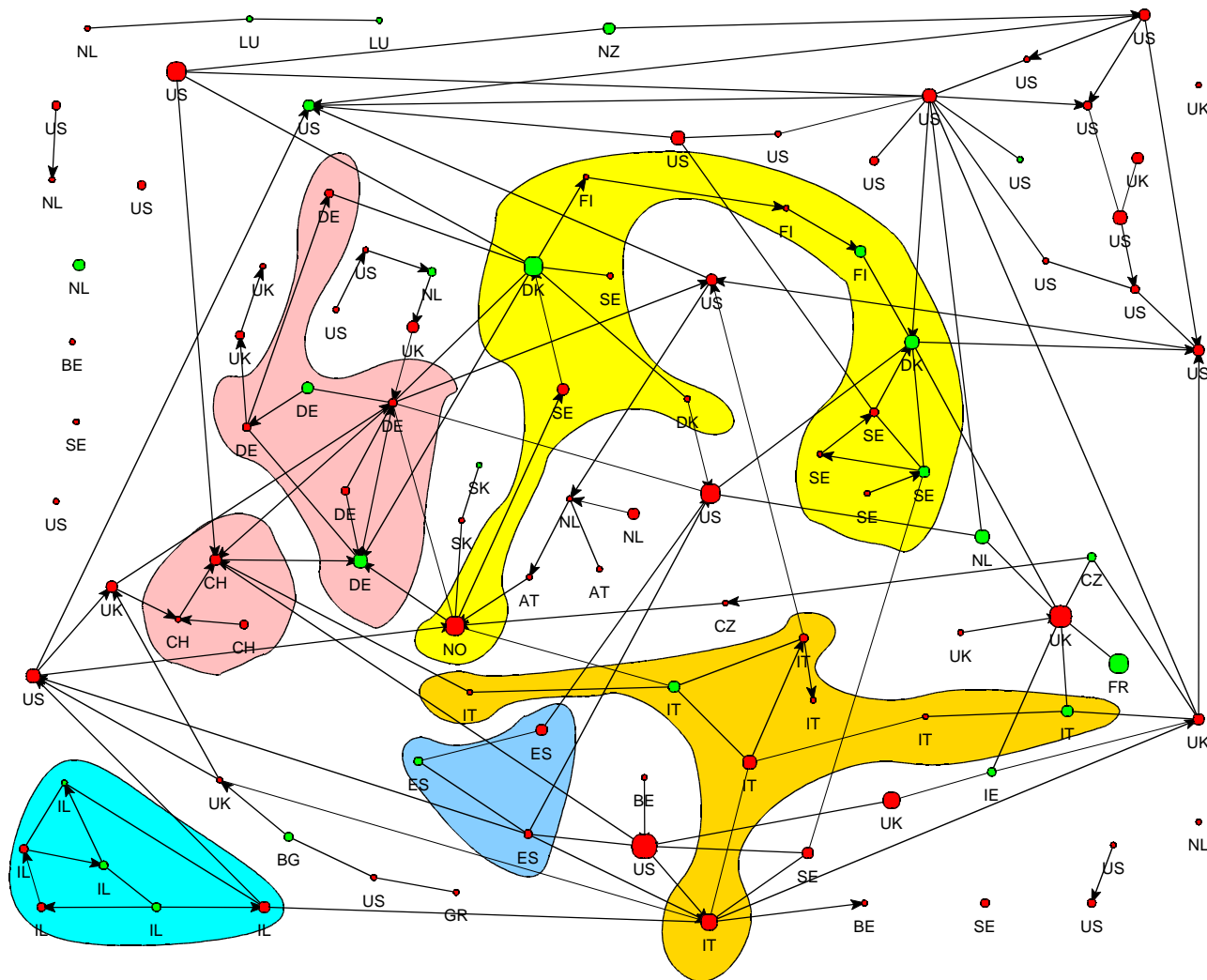


Figure 3: Small ASs in March 2001

From the three graphs of Figure 2, Figure 3 and Figure 4 several observations can be made:

- Only approximately half of the ASs observed have no test box. So packets sent by test boxes often cross ASs that also have a test box.
- Almost all countries observed have a test box (looking at Figure 1 for the large ASs and at Figure 2 for the small ASs, one can verify that in March 2000 only AT, BE and UK have no test box).
- The annealed graphs show remarkable stable geographical and cultural cohesion during three years. With this cohesion we mean that nodes in the graph that are close together in our drawings are also close together in geographical or cultural sense. This is a remarkable fact, because the proximity of nodes in our drawing is brought about by the annealing algorithm where mainly crossings are avoided, whereas geographical or cultural nearness means that ASs reside in the same country (IT, IL), are culturally close (Scandinavian countries) or are close together in the sense of language (DE, CH and AT, e.g.)

Scandinavian countries appear close together in all three annealed graphs. This also holds for German ASs and Italian ASs. Small Scandinavian ISPs always have at least one peering relation with another small Scandinavian ISP for three years in succession. Again the same holds for German and Italian ISPs.

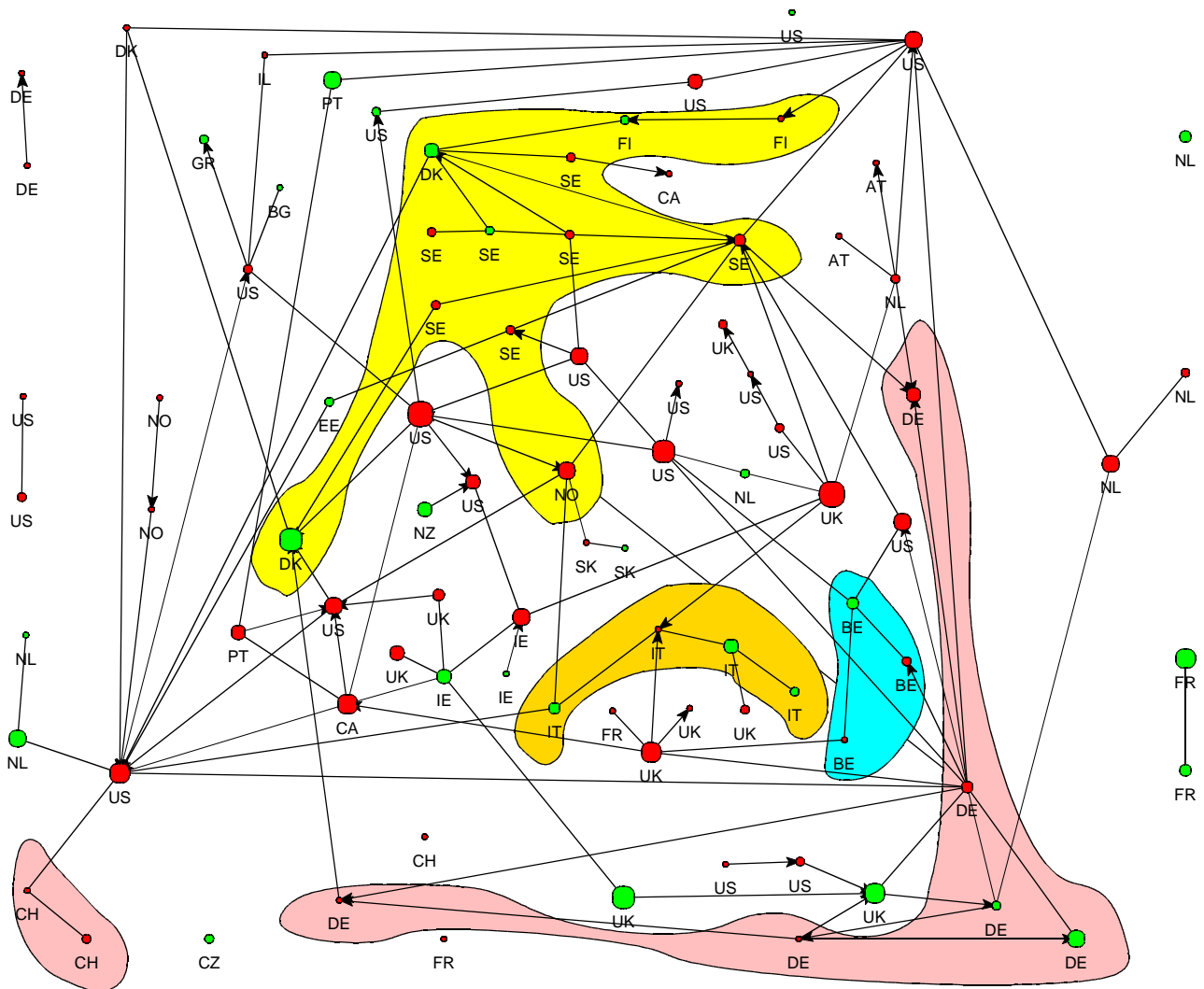


Figure 4: Small ASs in March 2002

- Geometrical cohesion does not hold for the many interspersed ASs that are located in the US. These ASs have head offices registered in the US, but we probably observe small subsidiaries in various European countries.
- Much of the observed traffic crosses ASs located in the US that do not have a test box. (from Figure 1 one can see that this also holds for the omitted large ASs). This means that ISP's (administratively) located in the US play an important role in the cohesion of the core the European part of the Internet.
- As an exception few peering agreements remain stable for three years where no geographical nearness exists. Examples are the configuration of a Norwegian AS with two Slovakian ASs (NO-SK-SK), or the one of two British and a Czech AS (UK-CZ-UK).

Table 4 summarises some numerical data of the three graphs in this paper. Also from this table some observations can be made:

- Only a very small fraction of the total ASs is observed by the TTM test boxes (1%, 120 of 12000). Because test boxes are placed close to border routers we mainly observe the core of the European part of the Internet.
- There are no big changes in three years. The number of observed IP-addresses, IP-edges and countries remains more or less constant. Also the degree distributions (see Figure 5) of the three AS-graphs are very similar. Thus the number of peering relations of ASs does not grow nor do ASs become bigger in the sense of observed IP addresses. This static impression might be caused by the fact that the number of active test boxes also did not grow (rather decreased a bit). In 2002 there is a decrease of the number of observed ASs compared to 2001, while the number of

active test boxes and observed IP-addresses remains constant. This may indicate that less providers become involved in the European core of the Internet.

- With the exception of the many ASs officially located in the US, the countries observed are mostly European, which can be expected considering that ISPs hosting a test box are frequently related to RIPE.

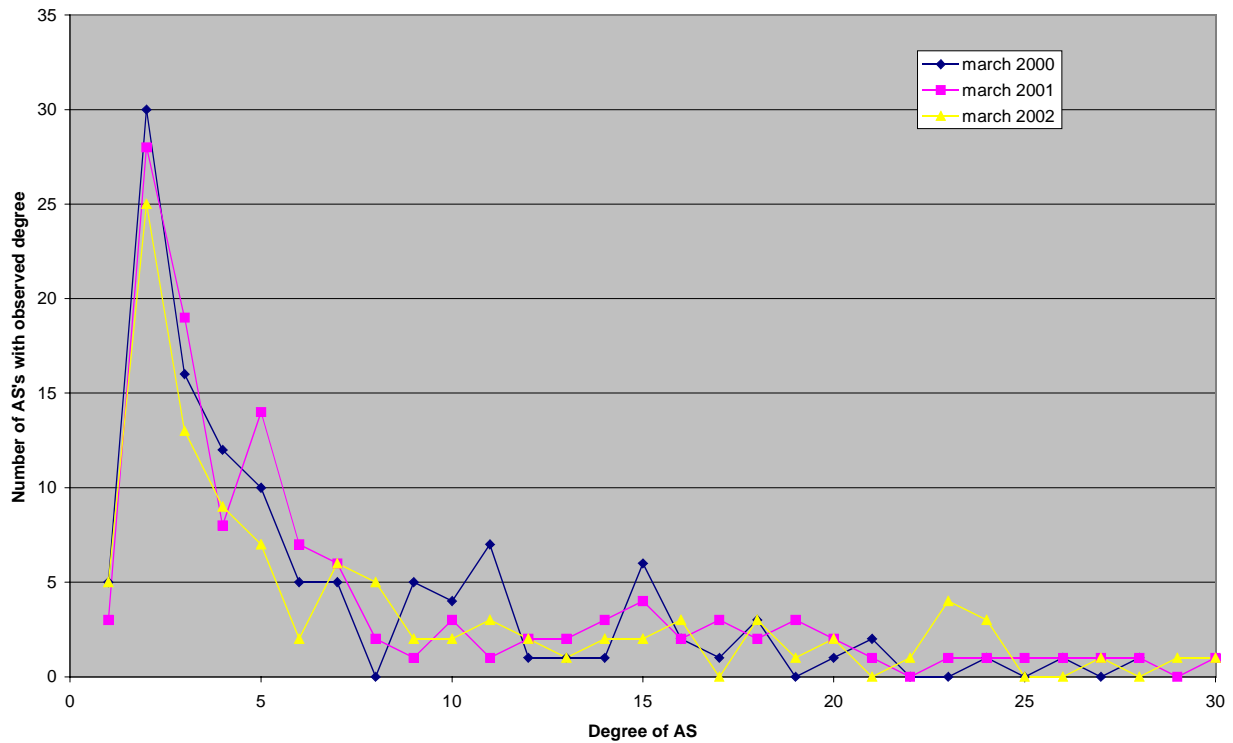


Figure 5: Degree distribution of the AS graphs of Figure 2, Figure 3 and Figure 4, but large ASs included.

Country	Small ASs				Large ASs				Total ASs			
	2000	2001	2002		2000	2001	2002		2000	2001	2002	
Austria	AT	1	2	2	AT				AT	1	2	2
Belgium	BE	2	3	3	BE				BE	2	3	3
Bulgaria	BG	1	1	1	BG				BG	1	1	1
Canada	CA			2	CA	1	1		CA	1	1	2
Switzerland	CH	7	3	3	CH	1	2	1	CH	8	5	4
Czech	CZ	1	2	1	CZ				CZ	1	2	1
Germany	DE	11	6	8	DE	1	2	1	DE	12	8	9
Denmark	DK	4	3	3	DK				DK	4	3	3
Estonia	EE			1	EE				EE			1
Spain	ES	3	3		ES				ES	3	3	
Finland	FI	4	3	2	FI				FI	4	3	2
France	FR	2	1	4	FR	1	1	1	FR	3	2	5
Greece	GR	5	1	1	GR				GR	5	1	1
Ireland	IE	3	1	3	IE				IE	3	1	3
Israel	IL	5	6	1	IL				IL	5	6	1
Italy	IT	7	8	4	IT				IT	7	8	4
Luxembourg	LU	2	2		LU				LU	2	2	
Netherlands	NL	4	8	7	NL	3	3	2	NL	7	11	9
Norway	NO	2	1	3	NO				NO	2	1	3
New Zealand	NZ		1	1	NZ				NZ		1	1
Poland	PT			2	PT				PT			2
Sweden	SE	13	9	7	SE	2	2	2	SE	15	11	9
Slovak	SK	2	2	2	SK				SK	2	2	2
United Kingdom	UK	6	11	9	UK	1	1	2	UK	7	12	11
United States	US	17	26	19	US	8	8	8	US	25	34	27
Total small		102	103	89	Total large	18	20	17	All ASs	120	123	106
% of all ASs		85,0	83,7	84,0	% of all ASs	15,0	16,3	16,0	# Test boxes	51	43	43
Degree		<15	<16	<17	Degree	≥15	≥16	≥17	# IP-nodes	3010	3235	3180
									# IP-edges	5533	5596	5360

Table 4: Small, large and total number of ASs in March 2000, 2001 and 2002

6. Discussion

Using topology data from the RIPE Test Traffic project we were able to show that the core of the European part of the Internet exhibits a remarkable geographical cohesion during the observation period of three years. We mapped IP addresses in the routing vectors to autonomous systems and then autonomous systems to countries. In the resulting graphs the geographical cohesion visually emerged when we used a simulated annealing algorithm to minimise the number of edge crossings.

In between the areas of geographical cohesion appear many interspersed autonomous systems of US origin, indicating that the overall cohesion of the core of the European part of the Internet depends on US originated ISPs.

In the three years of observation the size of the graph did not change much in terms of the number of countries, the number of IP addresses and the number of edges between IP addresses. In 2002, however, the number of observed autonomous systems decreased with 14%.

7. References

- [AKZ99] G. Almes, S. Kalidindi, M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.
- [CJW01] H. Chang, S. Jamin, W. Willinger, "Inferring AS-level Internet Topology from Router-Level Path Traces", SPIE ITCOM 2001, 2001
- [DH96] R. Davison, D. Harel, *Drawing Graphs Nicely Using Simulated Annealing*, ACM Transactions on Graphics, Vol. 15, No. 4, October 1996, Pages 301-331
- [HB96] J. Hawkinson, T. Bates, "Guidelines for creation, selection, and registration of an Autonomous System (AS)", RFC 1930, March 1996.
- [HPDC02] B. Huffaker, D. Plummer, D. Moore, and k. claffy, "Topology discovery by active probing", Symposium on Applications and the Internet (SAINT), Jan 2002.
- [IPPM] Internet Engineering Task Force, IP Performance Measurements Working Group (IETF IPPMWG), <http://www.ietf.org/html.charters/ippmcharter.html>
- [IRR] Internet Routing Registry, <http://www.irr.net/>
- [KZ99] Sunil Kalidindi and Matthew J. Zekauskas, "Surveyor: An Infrastructure for Internet Performance Measurements", INET'99 conference, 1999, <http://www.advanced.org/surveyor/>
- [MPDC00] D. Moore, R. Periakaruppan, J. Donohoe, and k. claffy, "Where in the world is netgeo.caida.org?", in INET '00, Yokohama, Japan, 18-21 Jul 2000, The Internet Society.
- [PAMM98] V. Paxson, G. Almes, J. Mahdavi, M. Mathis, "Framework for IP Performance Metrics", RFC 2330, May 1998.
- [RTT] RIPE Test Traffic working group home page: <http://www.ripe.net/ripe/wg/tt/index.html>
- [TRACES] Traces software package, <http://www.tbm.tudelft.nl/webstaf/janb/software/index.html>
- [UKKW98] Henk Uijterwaal, Daniel Karrenberg, Olaf Kolkman, René Wilhelm, "Internet Delay Measurements using Test traffic: First results", SANE98 Conference, Maastricht, the Netherlands, November 18-20, 1998