ISSN: 1467-1298

83
09 | 04

NET:GEOGRAPHY FIELDWORK FREQUENTLY ASKED QUESTIONS

Martin Dodge
Rob Kitchin

UCL

www.casa.ucl.ac.uk

T +44 (0) 20 7679 1782 • T +44 (0) 20 7679 1813 • F +44 (0) 20 7813 2843 • E casa@ucl.ac.uk

Centre for Advanced Spatial Analysis • University College London • 1 - 19 Torrington Place • Gower St • London • WC1E 7HB

NET:GEOGRAPHY FIELDWORK FREQUENTLY ASKED QUESTIONS

Martin Dodge

Centre for Advanced Spatial Analysis, University College London, 1-19 Torrington Place, Gower Street, London, WC1E 7HB, UK

Rob Kitchin

Department of Geography, National University of Ireland, Maynooth, Co. Kildare, Ireland.

Introducing the Net:Geography Fieldwork FAQ

Q. *What is the Net:Geography FAQ about?*

A. It is a set of answers to Frequently Asked Questions (FAQ) regarding the geography and unseen, 'inner' structures of the Internet. It provides a practical 'fieldwork' guide for understanding the Internet.

- It gives hands-on suggestions of techniques and freely available software tools and Web resources that can be used to actively explore both the internal topology of connection and the external geography of infrastructure. By revealing the operation of the Internet in terms of where things are located, who owns them and how data travels, the FAQ helps foster a more critical engagement with the media. The goal is to contribute, in a small way, to changing users of the Internet from passive consumers to more informed and active citizens of their network.
- It is possible to learn a lot about the Internet from critical writing, popular discourses and secondary published data. However, for real understanding, there is no substitute for doing your own fieldwork.
- It is not necessary to be a network engineer or computer scientist to begin to ask critical questions about the structure and operation of the Internet. Anyone can do some Net:Geography fieldwork.
- It does not require a large investment in expensive, specialised tools to undertake some Net:Geography fieldwork because the Internet can be used to measure and map itself. Many of the tools and techniques used in Net:Geography fieldwork were actually created by engineers for the practical purposes of 'debugging' network problem. However, they can also be re-used in politically challenging ways, providing tactical knowledge of the media that can not be gained in any other way.
- The practical examples given in this FAQ were tested using Windows 2000 PC on a university network, but most of the software tools and all of the techniques discussed are sufficiently generic that they should work in most situations. Although specific details may well vary, depending on the PC configuration and the type of connection to the Internet being employed.
- The Net:Geography FAQ comprises four sections: (1) finding out about your place on the Internet, (2) determining the location of components of the Internet (3) measuring distance across the Internet, (4) charting the routes of data through the Internet.

Q. *Can I really explore the 'inner workings' of the Internet without permission?*

A. Yes, even as an 'ordinary user' you can begin to explore the structure and operation of the Internet. This is because the Internet is built and operated in a fundamentally different way to other large communication networks, like telephones or television. These other networks are purposefully closed and proprietary, and, unlike the Internet, actively try to keep 'consumers' away from the insides of the network.

- The Internet was purposefully designed as an open network which encourages active exploration and experimentation. The Internet is not a single physical entity, instead it is premised on a public agreement to share data using open protocols. Anyone can use these protocols and as long as users abide by the terms of the agreement and follow the protocols, users are able to take an active role in producing the network. Many of the most useful Internet services widely used today came about through researchers, students and enthusiastic hackers exploring and exploiting this open architecture to try out new things.
- However, the openness of the Internet is always under attack because it is seen as threatening and subversive by many entrenched institutions. Today, with increasing commercial pressures, fears of criminal hackers, the floods of spam and waves of worms, there is definite 'chilling effect' across the Net. So don't be surprised if the active explorations of Net:Geography attracts the suspicions of 'officialdom'.
- For an lucid discussion of the design of the Internet, see Searls D. and Weinberger D. (2003) "World of ends: What the Internet is and how to stop mistaking it for something else" (www.worldofends.com).

Q. *Does the Internet actually have a geography?*

A. Yes it does. In fact there several different geographies, although this FAQ focuses on material geography of the infrastructures of the Internet.

- The hype around much of Internet, especially in the mid 1990s, was that it was 'everywhere and nowhere' and it would make geography less significant in human organisation through the 'death of distance'. This is patently not been the case.
- While the Internet has undoubtedly had an affect on the geography of business operation and individual consumption, distance is not dead. What is being witnessed is

a complicated restructuring, through processes of concentration and decentralisation, across scales

- The idea of the Internet as being somehow ‘anti-geographical’ is based on three key notions: fantasy, denial, and ignorance:

1. Internet geography was assumed not to exist. This is anti-corporeal, cyber-utopianist fantasy that somehow the virtual communities of cyberspace can be produced in a realm divorced from material existence.
2. Internet geography was assumed not to be important, so could be denied. The failure of many e-commerce ventures in the dotcom boom, we would argue, was based in part on ignoring the grounded, geographic, realities of computer-mediated communication, logistic networks and labour markets.
3. Internet geography was assumed not be measurable. Because it was hard to do, it was ignored, especially in the heady days of bubble growth.

- The medium of communication might be virtual, but the Internet is dependent on physical infrastructure and human labour, most of which is invisible to users. The computers are small in scale and are usually hidden from view in anonymous servers rooms and secure, windowless buildings, while the cables are under floors, in ceilings and in conduits buried under roads.
- The banal technicalities of Internet infrastructures are easily overlooked (just like for other essential utilities of water, electricity) but they are not naturally given. The geographical structure and operation of networks that service modern living have politics. Net:Geography fieldwork can help you grasp some of these politics first hand.
- Here are useful references which discuss the issues in more depth:
    - Castells M. (2001). *Internet Galaxy: Reflections on the Internet, Business and Society* (Oxford University Press, Oxford).
    - Cairncross F. (1997). *The Death of Distance* (Harvard Business School Press, Cambridge, MA).
    - Dodge M. and Kitchin R. (2001). *Mapping Cyberspace* (Routledge, London).
    - Graham S. and Marvin S. (1996). *Telecommunications and the City* (Routledge, London).

Q. *Why is understanding the geography of the Internet useful?*

A. There are several pragmatic reasons why knowing the geographical structure of the Internet is useful. Most importantly, the Internet is a global system, but is it always a locally produced. Understanding the local variability enhances understanding of the whole system.

- The social production of the Internet is contingent on cultural, legal, and economic forces that vary from place to place. In communicating with people it is often useful to be sensitive to language, customs and time-zones differences for example.
- The production of the Internet is subject to myriad of different legal systems, which vary by territorial geography. It can be important to know the legal jurisdiction where the user is located as this may impact the types of consumer protection enjoyed, the particular obscenity laws enforced, and so on.
- The freedom to surf the Web is not universal. Governments in many countries try to impose varying degrees censorship in the production and consumption of information of their citizens. (For an authoritative catalogue of government's censorship efforts across the world, see the Reporters without Borders report, "The Internet Under Surveillance: Obstacles to the Free Flow of Information Online", 2003, www.rsf.org.)
- In economic terms the Internet availability (as measured by access speeds, reliability and cost) remains uneven across space and across different social groups. This has been characterised, often overly simplistic, as the 'digital divide'.
- Knowing where things are located is also useful analytical because variations in spatial patterns can often give researchers an insight into underlying processes. Geographic location is also one of the most effective means of indexing Internet data, enabling linkages to be made to a vast array of existing secondary data, such as demographic statistics from censuses and surveys. Geography also provides a familiar frame for presenting data about the Internet, giving context and additional meaning to numbers. Cartography remains a powerful means of information presentation.
- Lastly, in a world of evermore information and services on the Internet, geography will prove to be an invaluable way of segmenting, filtering and prioritising people's attention. As a rule people tend to be more interested in information that is local to them, rather than things that are distant.

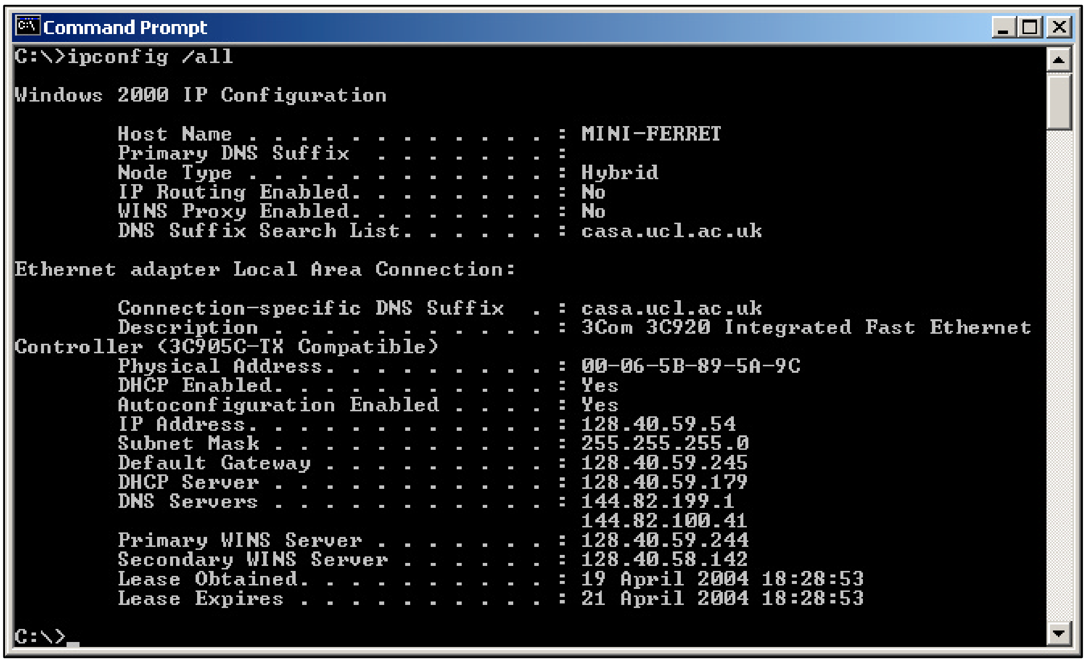
1. Finding out about your place on the Internet

We start the exploration of Net:Geography with some local fieldwork investigating how individuals are connected to the Internet and to see what is happening in their local Internet neighbourhood.

Q. *How am I connected to the Internet?*

A. Assuming you can see the physical equipment that you are using, what you want to know about is the more hidden parts of the connection in terms of software and the settings which identify you and your location to the rest of the Internet. It is quite easy to find out these details using diagnostic utilities of the operating system to display the current Internet configuration for your PC.

- Technically these are the TCP/IP (Transmission Control Protocol / Internet Protocol) settings. TCP/IP is the basic lingua franca of the Internet. If your computer is connected to the Internet it is ‘speaking’ in TCP/IP.
- The ipconfig utility will show the TCP/IP setting. Running it by typing ‘*ipconfig / all*’ from the command prompt gives the following type of output (Figure 1). (To open Command Prompt, click Start, then Programs, then Accessories.)



```
C:\>ipconfig /all

Windows 2000 IP Configuration

        Host Name . . . . . . . . . . . . : MINI-FERRET
        Primary DNS Suffix  . . . . . . . :
        Node Type . . . . . . . . . . . . : Hybrid
        IP Routing Enabled. . . . . . . . : No
        WINS Proxy Enabled. . . . . . . . : No
        DNS Suffix Search List. . . . . . : casa.ucl.ac.uk

Ethernet adapter Local Area Connection:

        Connection-specific DNS Suffix  . : casa.ucl.ac.uk
        Description . . . . . . . . . . . : 3Com 3C920 Integrated Fast Ethernet
Controller (3C905C-TX Compatible)
        Physical Address. . . . . . . . . : 00-06-5B-89-5A-9C
        DHCP Enabled. . . . . . . . . . . : Yes
        Autoconfiguration Enabled . . . . : Yes
        IP Address. . . . . . . . . . . . : 128.40.59.54
        Subnet Mask . . . . . . . . . . . : 255.255.255.0
        Default Gateway . . . . . . . . . : 128.40.59.245
        DHCP Server . . . . . . . . . . . : 128.40.59.179
        DNS Servers . . . . . . . . . . . : 144.82.199.1
                                            144.82.100.41
        Primary WINS Server . . . . . . . : 128.40.59.244
        Secondary WINS Server . . . . . . : 128.40.58.142
        Lease Obtained. . . . . . . . . . : 19 April 2004 18:28:53
        Lease Expires . . . . . . . . . . : 21 April 2004 18:28:53

C:\>_
```

Figure 1: TCP/IP settings reveal the structure of your local Net:Geography.

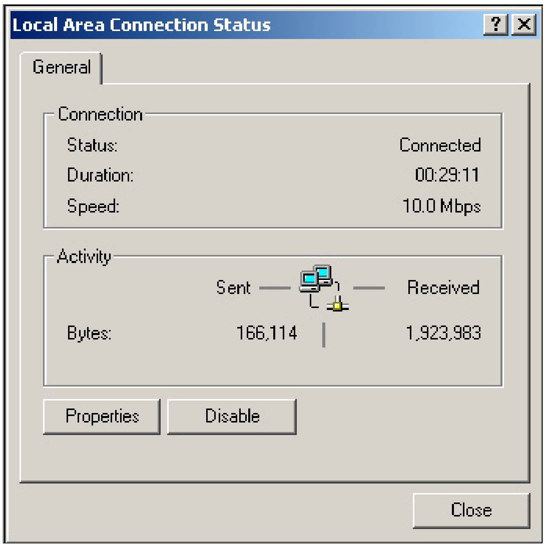
- The output looks technical (and it is to some extent) but all it shows is range of settings used to allow the PC online. The most useful parts to note are the name of your PC on the Internet (*Host Name: mini-ferret*) and the *IP Address (128.40.59.54*), which is the globally unique location of the PC in terms of the Internet's internal topology. No other computer on the Internet can (legally) share the same address.

- Details are also given on the type of connection, in this case through a local-area network using Ethernet, with the make and model of the card (*Description: 3Com 3C920*). The physical address of the card, a globally unique id code number that identifies this piece of hardware (*00-06-5B-89-5A-9C*) is also shown.
- Lastly, the output details the IP addresses of the *Default Gateway (128.40.59.245)*, which passes traffic from the local area out to the wider Internet, and the *DNS Servers (144.82.100.1 ; 144.82.100.41*), which are important components in the Internet for translating domain names into numeric IP addresses.

Q. *What is the speed of my connection?*

A. You can easily obtain the speed (and other useful statistics) on your current Internet connection.

- Open the Network and Dial-Up Connections menu (accessed from Start, Settings). Right-click on the active network connection and select the **status** option. You should see something like this.



```
Status:                Connected
Duration:               00:29:11
Speed:                 10.0 Mbps

Bytes:        166,114   1,923,983
```

Figure 2: Network status showing the speed of connection.

- The key component is obviously the speed. In this case the connection is running at *10 Mbps* (megabits per second). This is typical for an office environment and is quite a lot faster than average home Internet access. Also displayed is the duration of the session and a basic indication of activity in terms of the total data transferred in and out.

- The speed of connection is important because it determines the bandwidth you have for Internet interactions. Bandwidth is the capacity to shift data measured in bit per second and is crucial to what you can do and how long it will take, as illustrated by Figure 3. Surfing the Web via a dial-up telephone connection can seem very slow if you are use to faster Ethernet connection. Some Internet services are simply not viable over low bandwidth connections.

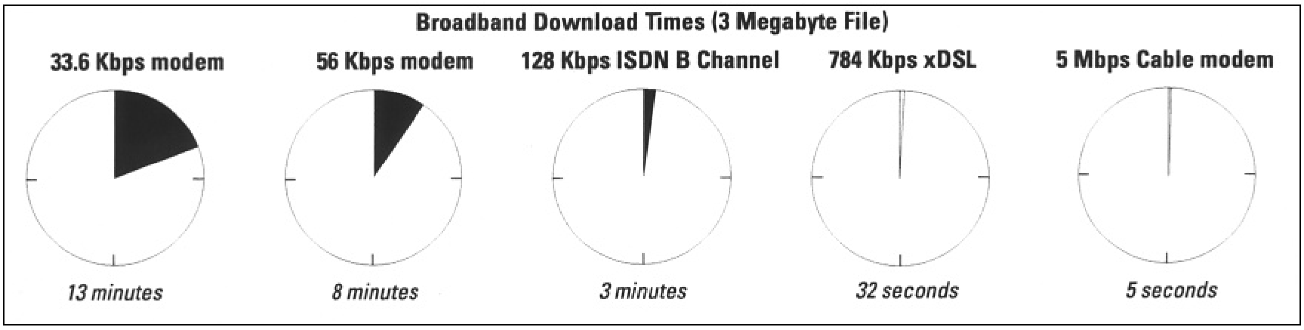
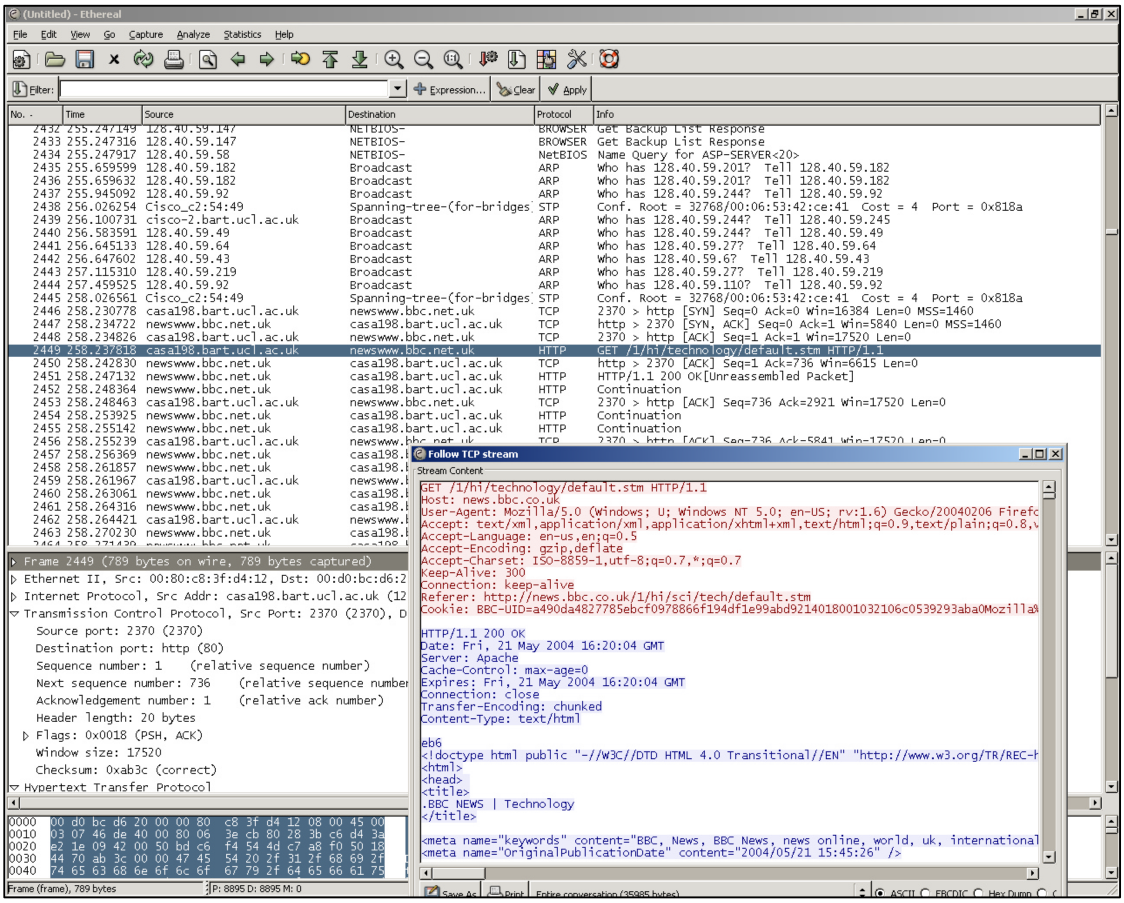


Figure 3: The importance of bandwidth (Source: TeleGeography, Washington D.C., www.telegeography.com).

Q. *What is going on in my local Internet neighbourhood?*

A. It is possible to see in some detail the inner structure of your local part of the Internet, observing the activities of other users, using network monitoring tools.

- There are lot of different tools available. One of the most capable is the Ethereal Network Analyzer. It is a free, opensource application, downloadable from *www.ethereal.com*.
- Ethereal is a network monitoring tool that does ‘packet sniffing’. This means it can watch all the traffic going around a local network. This traffic is in the form of streams of individual data packets flowing between different machines. The data packets can be captured by Ethereal for processing and detailed analysis.
- Figure 4 shows Ethereal monitoring a small local area network of an office at a university. It was able ‘sniff’ quite a lot of activity nearby. This screenshot only shows a snapshot of a few seconds of the data packets flowing past the monitoring PC.



```
No. -  Time       Source                    Destination                 Protocol  Info
2432 255.247149 128.40.59.147              NETBIOS-                    BROWSER   Get Backup List Response
2433 255.247316 128.40.59.147              NETBIOS-                    BROWSER   Get Backup List Response
2434 255.247917 128.40.59.58               NETBIOS-                    NetBIOS   Name Query for ASP-SERVER<20>
2435 255.659599 128.40.59.182              Broadcast                   ARP       Who has 128.40.59.201?  Tell 128.40.59.182
2436 255.659632 128.40.59.182              Broadcast                   ARP       Who has 128.40.59.201?  Tell 128.40.59.182
2437 255.945092 128.40.59.92               Broadcast                   ARP       Who has 128.40.59.244?  Tell 128.40.59.92
2438 256.026254 Cisco_c2:54:49             Spanning-tree-(for-bridges  STP       Conf. Root = 32768/00:06:53:42:ce:41  Cost = 4  Port = 0x818a
2439 256.100731 cisco-2.bart.ucl.ac.uk     Broadcast                   ARP       Who has 128.40.59.244?  Tell 128.40.59.245
2440 256.583591 128.40.59.49               Broadcast                   ARP       Who has 128.40.59.244?  Tell 128.40.59.49
2441 256.645133 128.40.59.64               Broadcast                   ARP       Who has 128.40.59.27?  Tell 128.40.59.64
2442 256.647602 128.40.59.43               Broadcast                   ARP       Who has 128.40.59.6?  Tell 128.40.59.43
2443 257.115310 128.40.59.219              Broadcast                   ARP       Who has 128.40.59.27?  Tell 128.40.59.219
2444 257.459525 128.40.59.92               Broadcast                   ARP       Who has 128.40.59.110?  Tell 128.40.59.92
2445 258.026561 Cisco_c2:54:49             Spanning-tree-(for-bridges  STP       Conf. Root = 32768/00:06:53:42:ce:41  Cost = 4  Port = 0x818a
2446 258.230778 casa198.bart.ucl.ac.uk     newswww.bbc.net.uk          TCP       2370 > http [SYN] Seq=0 Ack=0 Win=16384 Len=0 MSS=1460
2447 258.234722 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      TCP       http > 2370 [SYN, ACK] Seq=0 Ack=1 Win=5840 Len=0 MSS=1460
2448 258.234826 casa198.bart.ucl.ac.uk     newswww.bbc.net.uk          TCP       2370 > http [ACK] Seq=1 Ack=1 Win=17520 Len=0
2449 258.237818 casa198.bart.ucl.ac.uk     newswww.bbc.net.uk          HTTP      GET /1/hi/technology/default.stm HTTP/1.1
2450 258.242830 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      TCP       http > 2370 [ACK] Seq=1 Ack=736 Win=6615 Len=0
2451 258.247132 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      HTTP      HTTP/1.1 200 OK[Unreassembled Packet]
2452 258.248364 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      HTTP      Continuation
2453 258.248463 casa198.bart.ucl.ac.uk     newswww.bbc.net.uk          TCP       2370 > http [ACK] Seq=736 Ack=2921 Win=17520 Len=0
2454 258.253925 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      HTTP      Continuation
2455 258.255142 newswww.bbc.net.uk         casa198.bart.ucl.ac.uk      HTTP      Continuation
2456 258.255239 casa198.bart.ucl.ac.uk     newswww.bbc.net.uk          TCP       2370 > http [ACK] Seq=736 Ack=5841 Win=17520 Len=0
2457 258.256369 newswww.bbc.net.uk         casa198.b
2458 258.261857 newswww.bbc.net.uk         casa198.b
2459 258.261967 casa198.bart.ucl.ac.uk     newswww.b
2460 258.263061 newswww.bbc.net.uk         casa198.b
2461 258.264316 newswww.bbc.net.uk         casa198.b
2462 258.264421 casa198.bart.ucl.ac.uk     newswww.b
2463 258.270230 newswww.bbc.net.uk         casa198.b
```

```
▷ Frame 2449 (789 bytes on wire, 789 bytes captured)
▷ Ethernet II, Src: 00:80:c8:3f:d4:12, Dst: 00:d0:bc:d6:2
▷ Internet Protocol, Src Addr: casa198.bart.ucl.ac.uk (12
▽ Transmission Control Protocol, Src Port: 2370 (2370), D
    Source port: 2370 (2370)
    Destination port: http (80)
    Sequence number: 1    (relative sequence number)
    Next sequence number: 736    (relative sequence number
    Acknowledgement number: 1    (relative ack number)
    Header length: 20 bytes
  ▷ Flags: 0x0018 (PSH, ACK)
    Window size: 17520
    Checksum: 0xab3c (correct)
▽ Hypertext Transfer Protocol
```

```
0000  00 d0 bc d6 20 00 00 80  c8 3f d4 12 08 00 45 00
0010  03 07 46 de 40 00 80 06  3e cb 80 28 3b c6 d4 3a
0020  e2 1e 09 42 00 50 bd c6  f4 54 4d c7 a8 f0 50 18
0030  44 70 ab 3c 00 00 47 45  54 20 2f 31 2f 68 69 2f
0040  74 65 63 68 6e 6f 6c 6f  67 79 2f 64 65 66 61 75
```

Frame (frame), 789 bytes | P: 8895 D: 8895 M: 0

```
GET /1/hi/technology/default.stm HTTP/1.1
Host: news.bbc.co.uk
User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.0; en-US; rv:1.6) Gecko/20040206 Firefc
Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,v
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
Referer: http://news.bbc.co.uk/1/hi/sci/tech/default.stm
Cookie: BBC-UID=a490da4827785ebcf0978866f194df1e99abd9214018001032106c0539293aba0Mozilla%

HTTP/1.1 200 OK
Date: Fri, 21 May 2004 16:20:04 GMT
Server: Apache
Cache-Control: max-age=0
Expires: Fri, 21 May 2004 16:20:04 GMT
Connection: close
Transfer-Encoding: chunked
Content-Type: text/html

eb6
<!doctype html public "-//W3C//DTD HTML 4.0 Transitional//EN" "http://www.w3.org/TR/REC-h
<html>
<head>
<title>
.BBC NEWS | Technology
</title>

<meta name="keywords" content="BBC, News, BBC News, news online, world, uk, international
<meta name="OriginalPublicationDate" content="2004/05/21 15:45:26" />
```

Figure 4: Ethereal Network Analyzer capturing packet-level detail on the traffic flows of a the local network environment.

- The display looks complicated as it shows a very detailed view of Internet activity that most people never see. The result is a simply a long list of all traffic ‘sniffed’ rather than a summary graph or a map. Ethereal does not know anything about the physical structure of the network or the actual locations of the users. However, it is able to identify the different types traffic flowing and most importantly the source and destination of the traffic.
- The top window in the Ethereal interface displays one data packet per line. It takes some care and skill to interpret what is going in terms of user activities because they can be fragmented over many individual data packets. As an example, one particularly interesting data packet, number *2449*, has been selected. The data was sent from PC identified as *casa198.bart.ucl.ac.uk* (Source column) to *newswww.bcc.net.uk* (Destination column). The destination is the Web server for the BBC News service. The protocol of the data packet was *http* (used by Web browsers) and the *GET* command Info column reveals it was a request for a web page. Ethereal can connected

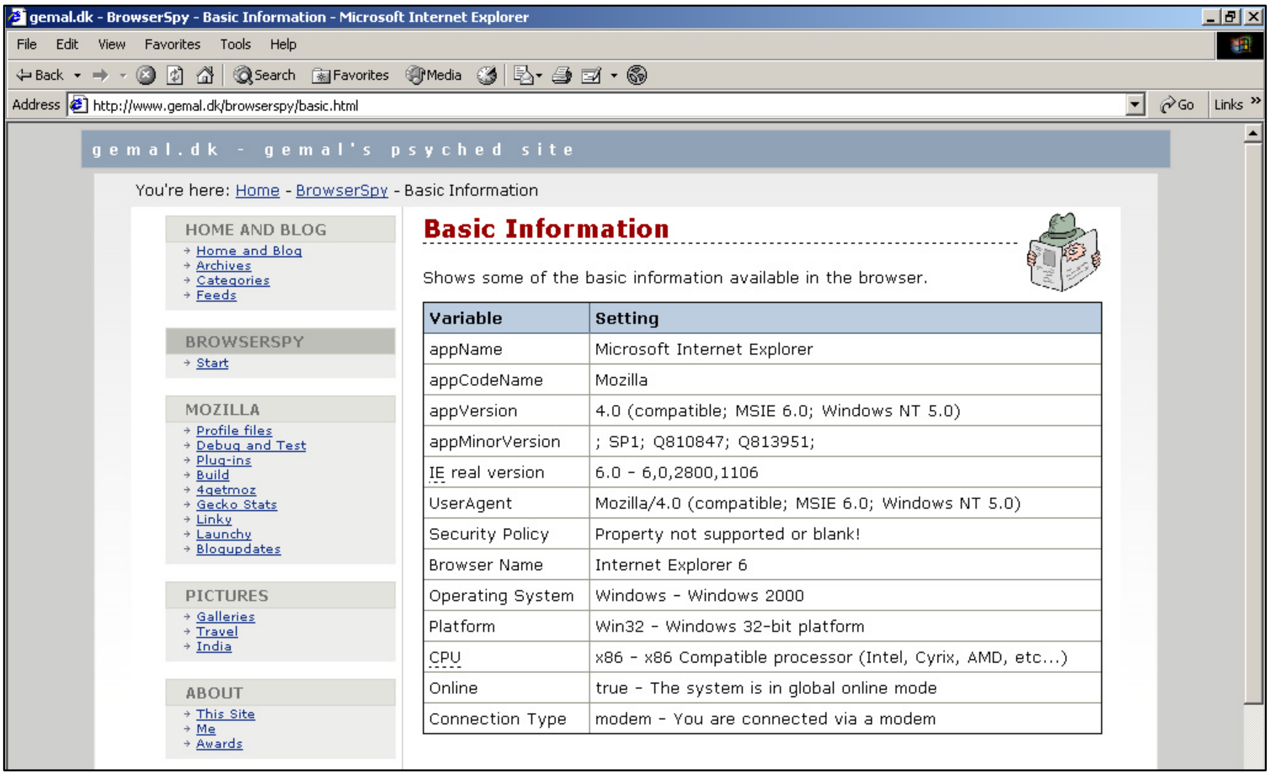
together all the related packets of data for this particular Web transaction, shown in the 'Follow TCP Stream' pop-up window, enabling us to see the actual Web page content (in raw html form). In this case Ethereal was able to show exactly the Web page this user was looking at.

- Ethereal, and similar network monitoring tools, have power to reveal a great deal about individual user's activities. This is particularly so because most data flowing across the Internet is sent unencrypted and can be covertly read by anyone able to tap into the network flow.
- Beyond the practicalities of running network monitoring tools like Ethereal and interpreting the output, there are clearly some more thorny ethical issues to confront about covert 'spying' on the activities of your neighbours. Good ethical practice for researchers would require that prior informed consent is obtained from of *all* people on the network being are scanned. This is hard to do. Active network scanning for benign Net:Geography exploration can also be deemed improper behaviour by strict system administrators, so you need to be prepared to justify your actions.

Q. *What do I reveal about myself to the rest of the Internet?*

A. Connecting to the Internet means necessarily revealing some details to network providers and leaving traces in logs in the services you interact with. At a most basic level the IP address must be known to send data to the correct location.

- Any online activities leave traces, but using different services and different client software results in different amounts of potentially personally identifiable data 'leaking' out.
- Surfing the Web in particular means you can be revealing a considerable amount of information without realising it. Even though you have not formally registered with websites and feel that you are browsing anonymously, you may be surprised the degree to which you are trackable.
- There are number of free web services that test what is revealed about your PC and Web browser configuration. Here is an example produced by the BrowserSpy service (http//gemal.dk/browserspy).


gemal.dk - BrowserSpy - Basic Information - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://www.gemal.dk/browserspy/basic.html
Go Links
gemal.dk - gemal's psyched site
You're here: Home - BrowserSpy - Basic Information
HOME AND BLOG
Home and Blog
Archives
Categories
Feeds
BROWSERSPY
Start
MOZILLA
Profile files
Debug and Test
Plug-ins
Build
4getmoz
Gecko Stats
Linky
Launchy
Blogupdates
PICTURES
Galleries
Travel
India
ABOUT
This Site
Me
Awards

Basic Information

Shows some of the basic information available in the browser.

Variable	Setting
appName	Microsoft Internet Explorer
appCodeName	Mozilla
appVersion	4.0 (compatible; MSIE 6.0; Windows NT 5.0)
appMinorVersion	; SP1; Q810847; Q813951;
IE real version	6.0 - 6,0,2800,1106
UserAgent	Mozilla/4.0 (compatible; MSIE 6.0; Windows NT 5.0)
Security Policy	Property not supported or blank!
Browser Name	Internet Explorer 6
Operating System	Windows - Windows 2000
Platform	Win32 - Windows 32-bit platform
CPU	x86 - x86 Compatible processor (Intel, Cyrix, AMD, etc...)
Online	true - The system is in global online mode
Connection Type	modem - You are connected via a modem

Figure 5: Some of the 'hidden' details that you can unwittingly reveal to websites.

- The 'basic information' that BrowserSpy is able to extract is perhaps not that suprising in that it knows the type of Web browser and version, as well as the operating system. More interestingly it could also tell the connection was via a modem. BrowserSpy is also able to gather many more details from the typical Web browser (e.g. plugins available, drive letters, screen resolution, time-zone settings). In many ways this is technical and very banal information, but taken together this voluntary 'leakage' can paint a detailed picture of your PC. This data is useful for websites in building profiles and tracking their users, and also for automatically presenting tailored content to suit different users (e.g. less graphics for those on low bandwidth connections). It can also be exploited by more unscrupulous people to identify potential vulnerabilities in your PCs through which to break in.
- Browser cookies are particularly pernicious tool in Web surveillance, although we do not have space to discuss their role in actively tracking individual browsing patterns. However, as a very simple test of their prevalence, try setting your browser so it has to request permission from you every time a website wants to set a cookie. You will quickly see just how many are set!

- Further technical reading on what data is known about you when using the Internet see, Richard C. (undated). “The Limits of Traceability” (www.cl.cam.ac.uk/users/rnc1/The_Limits_of_Traceability.pdf).

2. Location on the Internet

We now look beyond the local neighbourhood to some ways to explore Net:Geography more widely by looking at fieldwork activities in determining where things are located on the Internet.

Q. *How is location in the Internet defined?*

A. Knowing where things are located is crucial to most activities, including on the Internet. However, the Internet is concerned with topological location (the where things are located in terms of connections) and not with physical geographical location defined by x,y co-ordinates.

- At the foundation of the Internet is a robust and scaleable system to specify location uniquely so that data can be correctly transferred.
- As demonstrated earlier when discussing *ipconfig*, globally unique locations in the Internet are based on two key systems – IP addresses and domain names.
- Typically when people give out an Internet location this will be some kind of domain name address (e.g. a website address, *www.jasonnolan.net* or an email address such as *jason.nolan@utoronto.ca*). Domain names are always associated with an IP address.
- IP addresses are little seen and used by typical users. They look a little like telephone numbers. A typical example is *64.246.60.38* identifying *jasonnolan.net* website. The unique allocation of IP addresses are the key to the successful delivery of traffic across the complexity of the global Internet.

Q. *Why don’t Internet addresses specify a geographic position?*

A. This is how they were designed. Basically, the Internet protocols at the core of the network were never designed to make details on its geographical structure explicit to users. The successful transfer of data through the Internet requires knowledge of topological address, not geographical location.

- In many ways the Internet was designed to deliberately hide the underlying physical geography. This split between logical locations and physical locations is actually very useful. It is part of the reason why you can surf across websites scattered across the world and not have to worry about where the data is physically coming from.

- Because the Internet is a network of networks, rather than a homogeneous entity, its means of control are decentralised and its structure is fluid. Even though ‘no one owns the Internet’ as a whole, each one of its component parts is owned and operated by many millions of different organisations and individuals. In consequence, no one institution has a synoptic view of the whole Internet and, therefore, no one is required to register where all the components are physically located.
- However, it is often useful to be able to determine the geographic location of parts of the Internet, for example the location of a website as an part of assessment of whether they are credible sources of information or before giving them your credit card details. Many websites might not be physically where you think they are. However, because there is no one central register to do this mapping, you have to use a range of different fieldwork techniques which we run through below.

Q. *What types of geographic location are relevant for Net:Geography fieldwork?*
A. Understanding the nature of the geographical location of components of the Internet is interesting because different parts can be in different places. There are five obvious kinds of geographical location which are important in terms of the Web:

1. A website is where it is published, that is where the server is physically located (hardware geography).
2. A website is where the author/maintainer is located (production geography).
3. A website is where the legal owner is located (ownership geography).
4. A website is where the readers are located (audience geography).
5. A website is where the content refers (lexical geography).

- In some cases all five locations will be coincident. But it is easy to imagine plausible scenarios in which you have a Web page about Maynooth, Ireland that is written by someone in Canada, hosted on a server in London and read by people from across the world.
- The geographical precision of these different physical locations can also vary. Sometimes location might be determined as the precise x,y position (e.g. street address of the building with the web server) other times one might only know city or legal jurisdiction referred to.

Q. *How can I tell where an Internet address is geographically located?*

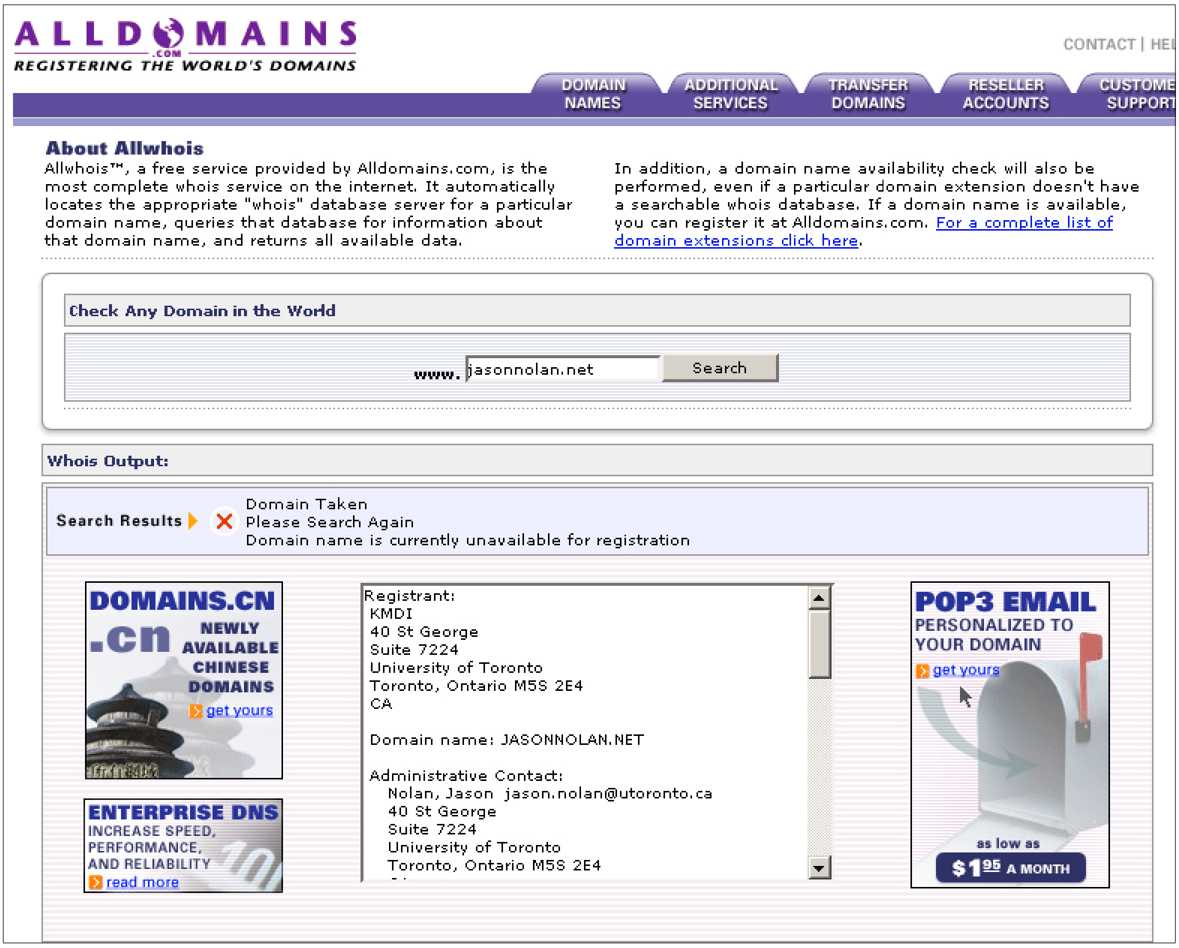
A. Unfortunately, determining the precise geographic location of an Internet address can not be done easily or in an accurate, consistent fashion. However, there are some techniques that can be used to try to determine the geographic location, at least approximately, of a website for example.

- The first point to note is that different techniques will tell you about the five different locational types. Most of the techniques described here will identify the production or ownership geography.
- The first, and most obvious, method is to use lexical location as the proxy. So you simply look at the content of the website to try to find an 'about page' or 'contacts page' that gives a postal address or telephone number of the owner. Other cultural (e.g. flags, symbols) and linguistic clues in the content of a website might give useful indications of the 'real-world' location. Ultimately, of course, you could try asking them directly where they are located assuming they provide a contact email or telephone number.
- If you only have the domain name of a website to work with, the first place to start is by 'decoding' this. Many domain names are allocated on a country by country basis with their name ending with the appropriate two letter ISO country code (e.g. *.ca* for Canadian domains, *.ie* for Irish domains and so on). One can infer the geographic location of an Internet address based on the country code in its domain name. A useful list of all the country code domain names is available at *www.iana.org/cctld/cctld-whois.htm*.
- However, there are limitations in relying on country-code domains.
    1. There are several domain name types (so called 'global top-level domains') that are not related to countries. The biggest of these are *.com, .org, .net* and *.info*. The *jasonnolan.net* website has a *.net* domain name which does not allow any inference on its location to be made. A *.net* domain could be located anywhere in the world. (Note, the top-level domains *.mil, .edu* and *.gov* are allocated only to U.S. institutions, so one can be fairly safely assume they are located in the USA.)
    2. The level of geographic precision is obviously crude with this technique, particularly so for large countries. A *.ca* domain name could be anywhere in Canada.

3. Lastly, just because a website uses a country code domain name does not guarantee that the website is actually within the country indicated. The ownership, production, hosting and use of that website could well be in another country or several different countries. For example, the amazon.de and amazon.co.uk websites are running on servers physically located in the U.S and not in the United Kingdom or Germany as indicated by their domain names.

- Moving beyond decoding the top-level domain names, you can try to exploit other parts of the domain name to infer geographic location of the address. This works when the email addresses or websites are part of an established, easily identifiable company or institution that can be traced to city in which they are located. For example, we could deduce that Martin Dodge's email address at ucl.ac.uk links him to University College London (UCL) and could then infer that he is physically located in central London, on the site of the main university campus. Again, there are limits to the level of precision in this location in that it can not be determined, just from the ucl.ac.uk domain, what the actual street address is for Martin Dodge. This method is also error prone for large organisations, like transnational companies, which operate from many sites over extensive areas.
- The last thing you can do with a domain name address is to find out who it is registered to. All domain names have a legal owner and the registration databases for this usually contain contact details. You can freely consult this registration information from the domain name system using a *Whois* query. (Note, not all domain registration databases will publicly give out the full address details of the owner.)

- A *Whois* query can be easily run from any number of websites. Below is the results of a *Whois* query to find the registered owner of *jasonnolan.net* domain name using the free service provided by AllWhois (www.allwhois.com).



About Allwhois

Allwhois™, a free service provided by Alldomains.com, is the most complete whois service on the internet. It automatically locates the appropriate "whois" database server for a particular domain name, queries that database for information about that domain name, and returns all available data.

In addition, a domain name availability check will also be performed, even if a particular domain extension doesn't have a searchable whois database. If a domain name is available, you can register it at Alldomains.com. For a complete list of domain extensions click here.

Domain Taken
Please Search Again
Domain name is currently unavailable for registration

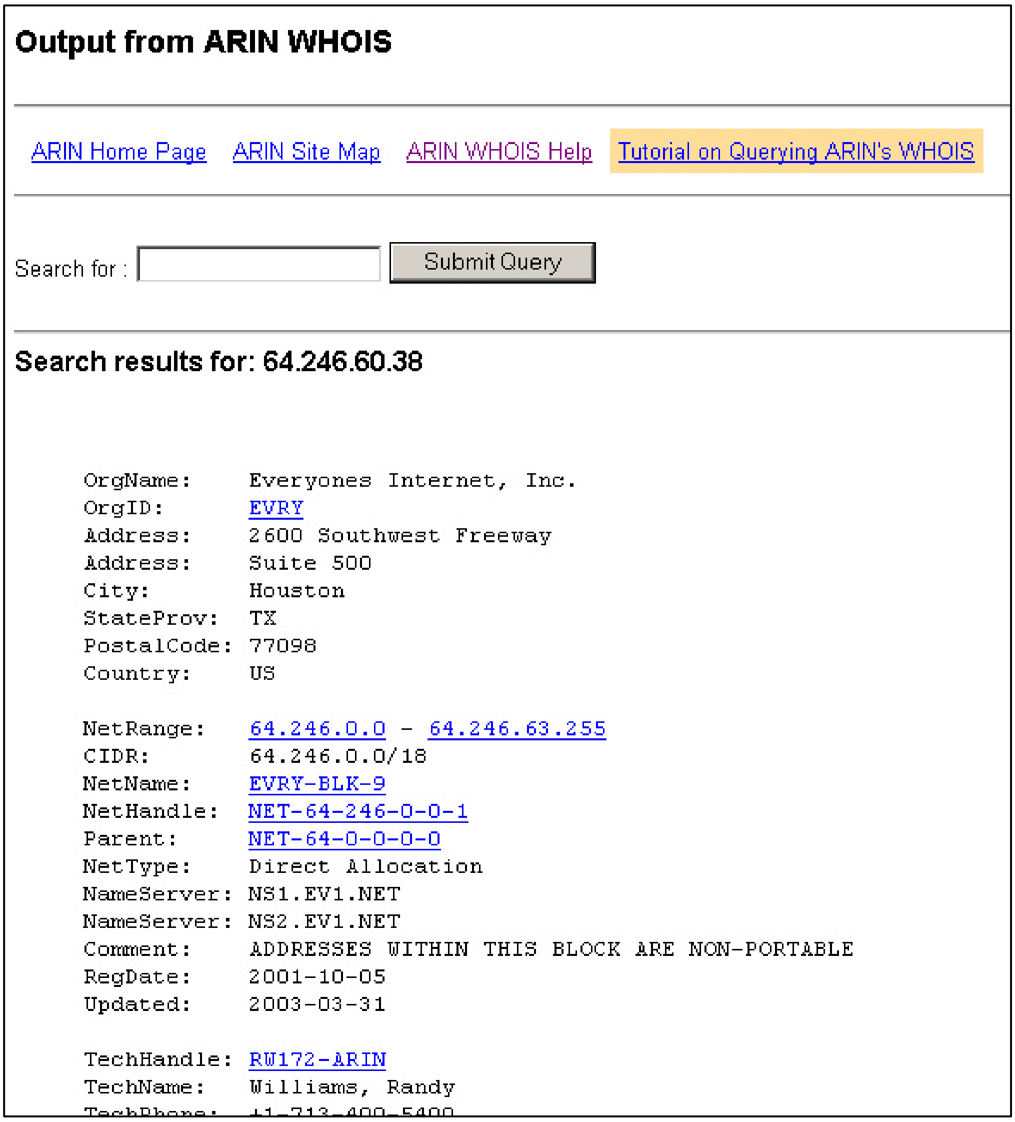
```
Registrant:
 KMDI
 40 St George
 Suite 7224
 University of Toronto
 Toronto, Ontario M5S 2E4
 CA

 Domain name: JASONNOLAN.NET

 Administrative Contact:
    Nolan, Jason  jason.nolan@utoronto.ca
    40 St George
    Suite 7224
    University of Toronto
    Toronto, Ontario M5S 2E4
```

Figure 6: Looking up the domain name registrations details on jasonnolan.net using Whois.

- The registration information from the *Whois* query identifies the owner of *jasonnolan.net* to be located in Toronto, Canada. The full street address is given. This could be looked up and a detailed map obtained giving the precise location.
- The results of *Whois* queries can be very useful in finding out the where the registered owner of a domain name is, however they are not always accurate. Firstly, registration details held on a given domain name may be out of date, incorrect or deliberately false (spammers, for example, try to hide their true geographic location and would be unlikely to complete the registration honestly). Secondly, ownership details may only tell you one of the five possible geographies of a website. We can see that *jasonnolan.net's* owner is located in Toronto, but the site may be produced, published, and consumed elsewhere. Thirdly, the registrations for large organisations often give a

single postal address of the headquarters and thereby mask where individual domain names are actually being used.

- If you do not have a domain name to work with, you can also lookup the ownership of IP addresses. These are generally allocated in large blocks to ISPs rather than to individuals or companies. You can do this query by doing a *Whois* query to ARIN (www.arin.net/whois). The IP address of the server which publishes *www.jasonnolan.net* is *64.246.60.38*. Looking up this address yields details on the registered owner, a company called Everyone's Internet Inc., with a postal address in Houston Texas.



Output from ARIN WHOIS

Search results for: 64.246.60.38

```
OrgName:    Everyones Internet, Inc.
OrgID:      EVRY
Address:    2600 Southwest Freeway
Address:    Suite 500
City:       Houston
StateProv:  TX
PostalCode: 77098
Country:    US

NetRange:   64.246.0.0 - 64.246.63.255
CIDR:       64.246.0.0/18
NetName:    EVRY-BLK-9
NetHandle:  NET-64-246-0-0-1
Parent:     NET-64-0-0-0-0
NetType:    Direct Allocation
NameServer: NS1.EV1.NET
NameServer: NS2.EV1.NET
Comment:    ADDRESSES WITHIN THIS BLOCK ARE NON-PORTABLE
RegDate:    2001-10-05
Updated:    2003-03-31

TechHandle: RW172-ARIN
TechName:   Williams, Randy
TechPhone:  +1-713-400-5400
```

Figure 7: The results of a Whois lookup on the IP address of the Web servers that hosts jasonnolan.net.

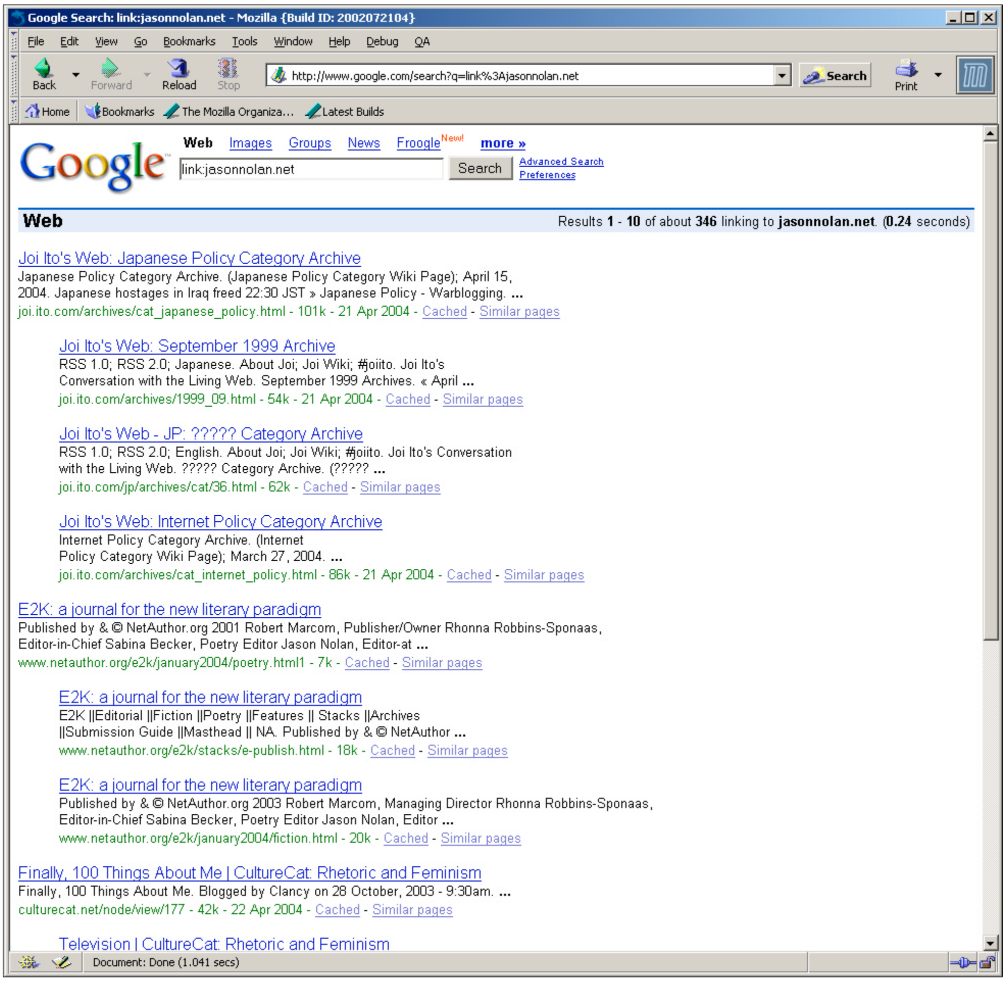
- Another thing you can do is actually test where a website is connected to the Internet by tracing traffic flows as this can give some useful clues to its physical location. Details on how to do this real-time probing are discussed in the last section of the FAQ.
- There are a number of firms that provide commercial services to convert IP addresses to geographic locations. For example, Quova, Inc (www.quova.com), IP2Location (www.ip2location.com). However, these services not really aimed at individual Net:Geography explorers.
- To find out more on the technicalities of relating Internet addresses to geographic locations these two computer science papers are useful places to start:
    - Lakhina, A., Byers, J.W., Crovella, M. and Matta, I., (2002). "On the Geographic Location of Internet Resources". *Technical Report 2002-15*, Computer Science Department, Boston University. (www.cs.bu.edu/techreports/pdf/2002-015-internet-geography.pdf)
    - Padmanabhan, V.N. and Subramaniann L., (2001). "Investigation of Geographic Mapping Techniques for Internet Hosts". *SIGCOMM'01*, August 27-31, San Diego. (www.research.microsoft.com/~padmanab/papers/sigcomm2001.pdf)

Q. *What else can I do to find out more about a website's ownership and location?*

A. Taking a different perspective, you can also explore the 'virtual' position of a website in terms of its visibility in the information space of the Web.

- Counting the number of hyperlinks to and from a website and analysing whom the links come from can reveal the informational structures of Net:Geography. The results can be used to infer a website's position in terms of social networks and power geometries; a well linked site could indicate that its creator has power and influence. This kind of hyperlink analysis has many parallels to citation analysis used to assess influence in scholarly research. Much of the success of the Google search engine in terms of relevance ranking depends on its analysis of hyperlink structures to indicate the most credible sources of information.
- It is possible to obtain appropriate data to give a rough approximation of hyperlink structures using some of the large Web search engines. These will report link statistics (usually available as part of their advanced search options). For example using Google

you can find out the number and origin of incoming hyperlinks made to the *jasonnolan.net* website using the search command *link:jasonnolan.net.*



Google Search: link:jasonnolan.net - Mozilla {Build ID: 2002072104}
File Edit View Go Bookmarks Tools Window Help Debug QA
Back Forward Reload Stop
http://www.google.com/search?q=link%3Ajasonnolan.net
Search
Print
Home Bookmarks The Mozilla Organiza... Latest Builds
Google
Web Images Groups News Froogle New! more »
Search
Advanced Search
Preferences

link:jasonnolan.net

Web Results **1** - **10** of about **346** linking to **jasonnolan.net**. (**0.24** seconds)

Joi Ito's Web: Japanese Policy Category Archive
Japanese Policy Category Archive. (Japanese Policy Category Wiki Page); April 15, 2004. Japanese hostages in Iraq freed 22:30 JST » Japanese Policy - Warblogging. **...**
joi.ito.com/archives/cat_japanese_policy.html - 101k - 21 Apr 2004 - Cached - Similar pages

Joi Ito's Web: September 1999 Archive
RSS 1.0; RSS 2.0; Japanese. About Joi; Joi Wiki; #joiito. Joi Ito's Conversation with the Living Web. September 1999 Archives. « April **...**
joi.ito.com/archives/1999_09.html - 54k - 21 Apr 2004 - Cached - Similar pages

Joi Ito's Web - JP: ????? Category Archive
RSS 1.0; RSS 2.0; English. About Joi; Joi Wiki; #joiito. Joi Ito's Conversation with the Living Web. ????? Category Archive. (????? **...**
joi.ito.com/jp/archives/cat/36.html - 62k - Cached - Similar pages

Joi Ito's Web: Internet Policy Category Archive
Internet Policy Category Archive. (Internet Policy Category Wiki Page); March 27, 2004. **...**
joi.ito.com/archives/cat_internet_policy.html - 86k - 21 Apr 2004 - Cached - Similar pages

E2K: a journal for the new literary paradigm
Published by & © NetAuthor.org 2001 Robert Marcom, Publisher/Owner Rhonna Robbins-Sponaas, Editor-in-Chief Sabina Becker, Poetry Editor Jason Nolan, Editor-at **...**
www.netauthor.org/e2k/january2004/poetry.html1 - 7k - Cached - Similar pages

E2K: a journal for the new literary paradigm
E2K ||Editorial ||Fiction ||Poetry ||Features || Stacks ||Archives ||Submission Guide ||Masthead || NA. Published by & © NetAuthor **...**
www.netauthor.org/e2k/stacks/e-publish.html - 18k - Cached - Similar pages

E2K: a journal for the new literary paradigm
Published by & © NetAuthor.org 2003 Robert Marcom, Managing Director Rhonna Robbins-Sponaas, Editor-in-Chief Sabina Becker, Poetry Editor Jason Nolan, Editor **...**
www.netauthor.org/e2k/january2004/fiction.html - 20k - Cached - Similar pages

Finally, 100 Things About Me | CultureCat: Rhetoric and Feminism
Finally, 100 Things About Me. Blogged by Clancy on 28 October, 2003 - 9:30am. **...**
culturecat.net/node/view/177 - 42k - 22 Apr 2004 - Cached - Similar pages

Television | CultureCat: Rhetoric and Feminism

Document: Done (1.041 secs)

Figure 8: Incoming hyperlinks to jasonnolan.net website according to the Google index.

- The Google search engine index reports 346 web pages with hyperinks to *jasonnolan.net*. This is a respectable number of links. From a limited understanding, many of them appear to come from blogs citations. This is perhaps not surprising at the jasonnolan.net is also a blog.

- These types of informational structures can also viewed spatially as graphs, where the Web pages are nodes and the hyperlinks are connecting lines. A nice example of this is available using *GoogleBrowser*, an interactive tool produced by TouchGraph (www.touchgraph.com). It allows you to actively explore a website's location in terms of the virtual economy of linkages. Below is an example of the GoogleBrowser view of linkage network immediately around *www.jasonnolan.net*.

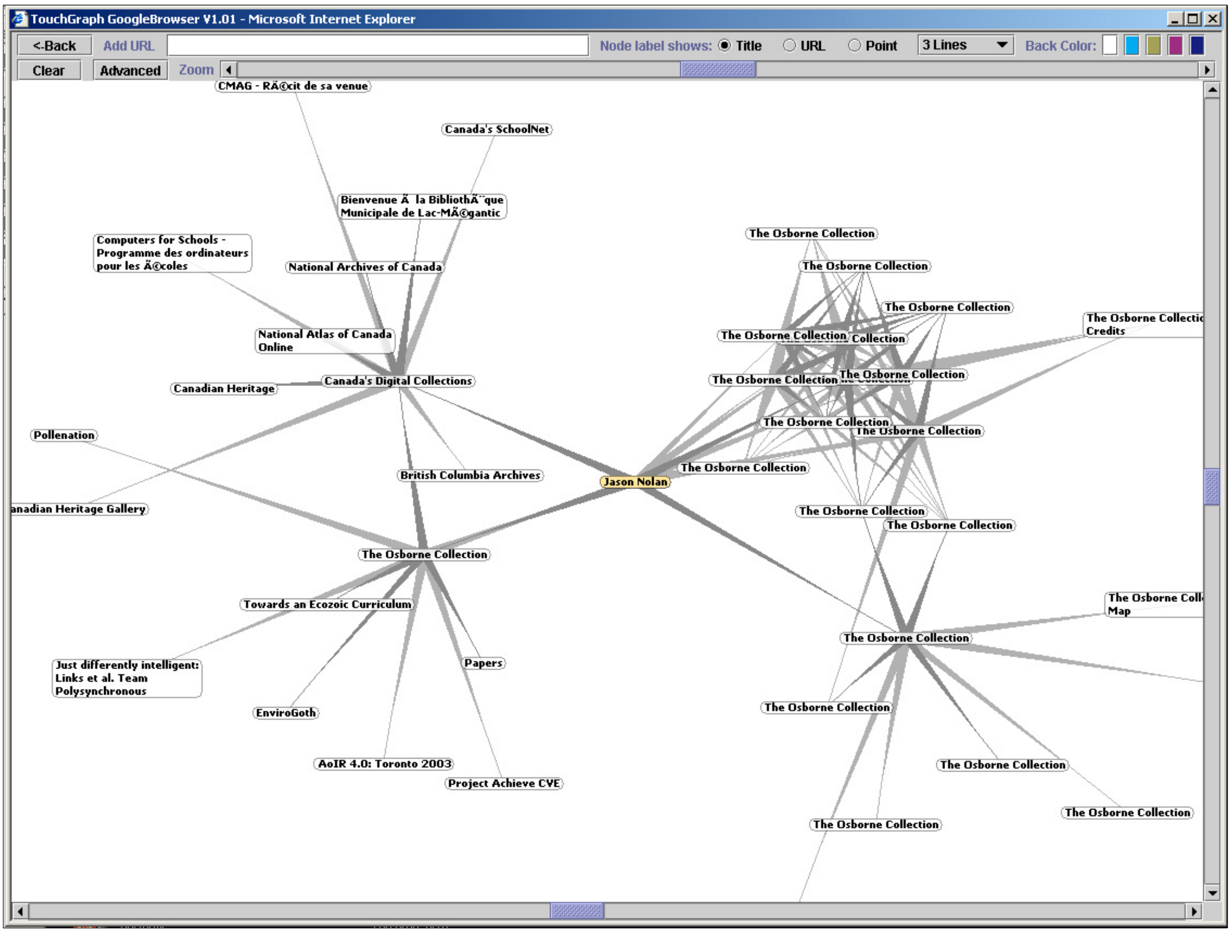

TouchGraph GoogleBrowser V1.01 - Microsoft Internet Explorer
<-Back
Add URL
Node label shows: ● Title ○ URL ○ Point
3 Lines
Back Color:
Clear
Advanced
Zoom
CMAG - RÃ©cit de sa venue
Canada's SchoolNet
Bienvenue Ã la BibliothÃ¨que Municipale de Lac-MÃ©gantic
Computers for Schools - Programme des ordinateurs pour les Ã©coles
National Archives of Canada
National Atlas of Canada Online
Canadian Heritage
Canada's Digital Collections
Pollenation
anadian Heritage Gallery
British Columbia Archives
Jason Nolan
The Osborne Collection
Towards an Ecozoic Curriculum
Just differently intelligent: Links et al. Team Polysynchronous
Papers
EnviroGoth
AoIR 4.0: Toronto 2003
Project Achieve CVE
The Osborne Collection
The Osborne Collectic Credits
The Osborne Coll Map
Collection

Figure 9: GoogleBrowser view of the local Web neighbourhood of jasonnolan.net.

- For those interested in exploiting the tactical power of hyperlink analysis to expose the hidden politics of Net:Geography, the research of Richard Rogers and colleagues at *www.govcom.org* is worth consulting.

3. Measuring distance across the Internet

Q. *How is distance across the Internet defined?*

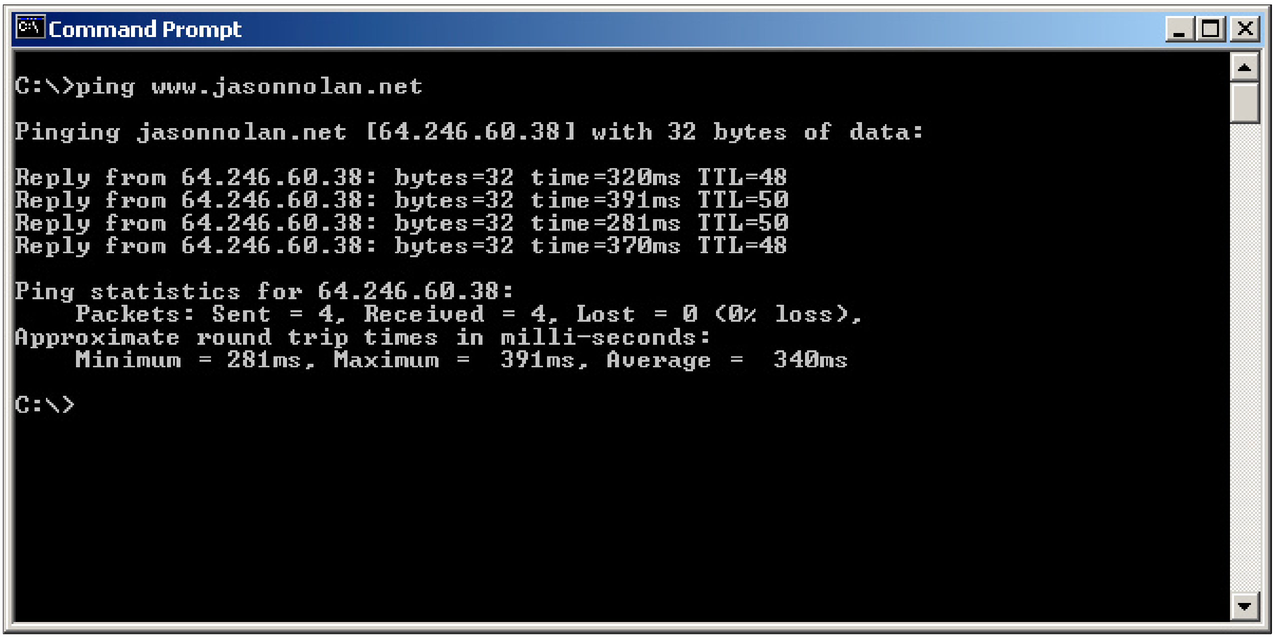
A. Simply put, distance is measured in terms of time inside the Internet.

- Within the topological structure of the Internet physical distances between places have little meaning or relevance. Instead relative distances are measured on the ‘journey’ time taken to transmit and receive data.
- This ‘journey’ time is called latency. Increasing latency implies increasing relative distance between two places on the Internet. Latency metrics serve the same purpose as physical distance on road signs and maps, that is they tell you how ‘near’ or ‘far’ apart things are.
- Of course time, and associated costs, are also widely used in the ‘real’ world for expressing distance. Most people assess a journey in terms of the time it takes and not linear distance on the ground (e.g. a 10 minute drive to the shops, a 4 hour flight to a holiday destination). Time-based distance measures are useful to people because they match subjective perceptions of closeness, relevance and importance. Near things tend to be judged as more important because people’s exposure to them is more direct, immediate and frequent. Far away things take longer to experience and thus tend to be less familiar (e.g. people speak a different language) and are perceived as less comfortable and perhaps even as more risky/dangerous.
- An interesting point of analysis, both on the Internet and in the ‘real’ world, is to compute the relationship between distance on the ground and time distance for different places. This relationship is not always linear because of barriers, lack of connectivity and poor accessibility. Sometimes the quickest places to reach are not always the closest physically. Analysing the variable patterns in time accessibility can sometimes give insight into the underlying structural processes.
- Distance can also be measured by the complexity of the journey, such as the number of interchanges required. In terms of the Internet, the least-complex distance would seek to minimise the number of different networks crossed or switching points negotiated, which may or may not be the same as the quickest route. The next section of the FAQ shows how you can plot the route complexity by tracing traffic flows.

Q. *How can I actually measure latency?*

A. There is a useful utility called *ping* that can measure latency (the travel time of data) on the Internet. Ping is easy to use and available on most PCs.

- It is a network measurement tool primarily used by engineers to diagnose connectivity problems. Basically, it reports whether the place on the Internet that is trying to be reach is ‘alive’.
- Ping takes its name from the sound that submarine sonar uses. Conceptually it works in a similar fashion by sending out small packets of test data to a target host and listening for a response. It is useful for distance measurement because it reports the round-trip time of the data packets.
- Here is an example of measuring the distance from a PC in London, UK to *www.jasonnolan.net* using ping.



```
C:\>ping www.jasonnolan.net

Pinging jasonnolan.net [64.246.60.38] with 32 bytes of data:

Reply from 64.246.60.38: bytes=32 time=320ms TTL=48
Reply from 64.246.60.38: bytes=32 time=391ms TTL=50
Reply from 64.246.60.38: bytes=32 time=281ms TTL=50
Reply from 64.246.60.38: bytes=32 time=370ms TTL=48

Ping statistics for 64.246.60.38:
    Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
    Minimum = 281ms, Maximum =  391ms, Average =  340ms

C:\>
```

Figure 10: Using ping to measure time distance between two points on the Internet.

- By default on Windows, ping sends out four test data packets. The time each took to go from London to *jasonnolan.net’s* server and back again is reported (in milliseconds). The last line of the output reports the overall statistics. According to this, the average ‘distance’ for this particular journey across the Internet as measured by latency was 340ms, while the slowest data packet took 391ms.
- This type of time distance measurement is very susceptible to changes in conditions on the Internet, especially congestion in traffic flows. Internet distances measured by

latency are never fixed. However, this variability is actually useful as it can be used as a way of quantifying the fluctuations in Internet conditions, much like measuring car speeds gives an indication of the level of road congestion.

Q. *What else can I do with ping?*

A. There are several ways that 'pinging' Internet distances can be used to learn more about Net:Geography.

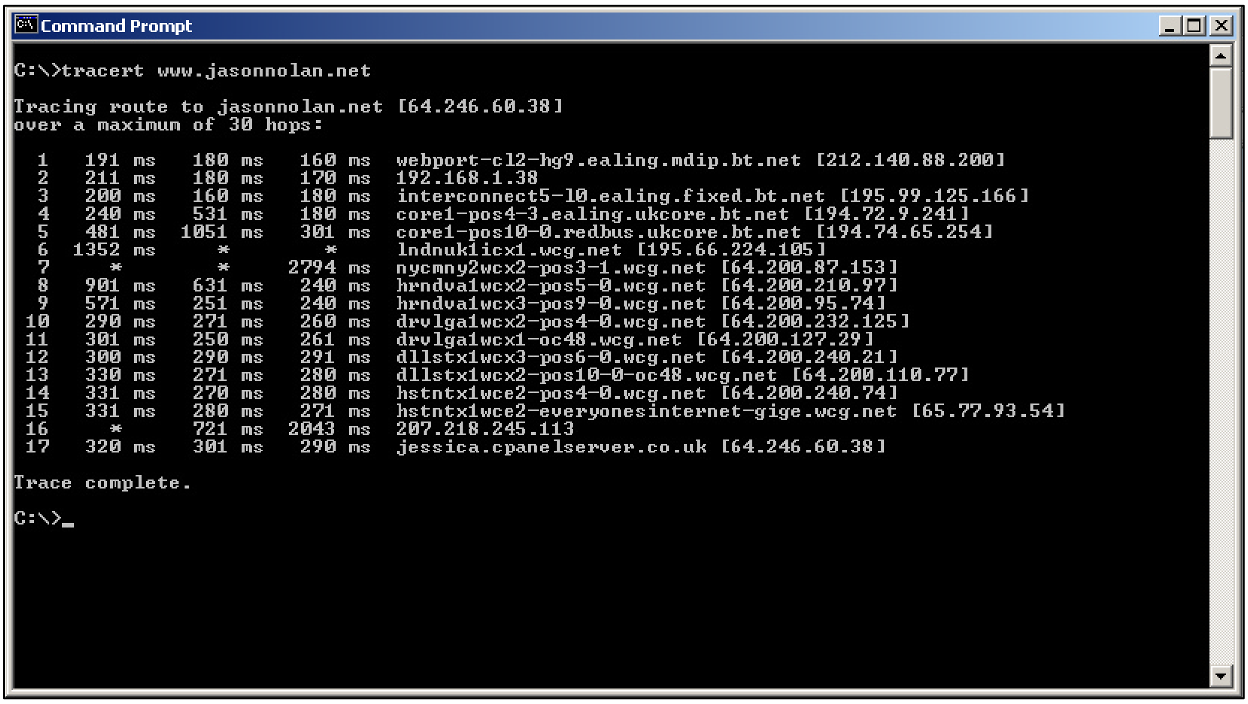
- First, and most obviously, a sequence of pings to the same place at different time periods can be used to build up a comprehensive longitudinal profile of latency. This might reveal interesting temporal patterns in the variation of latency as places on the Internet move nearer and further apart in a predictable fashion. Sudden changes in latency that do not fit the profile can reveal serious problems (e.g. a physical cut in a key fibreoptic cable).
- Another useful extension is to take pings from different places on the Internet to triangulate in on a particular target point. For example, one could take measurements of latency to *jasonnolan.net* not just from London, but from other geographically closer or and more distant points. The practicalities of doing this type of triangulation are made easier because there are a number of websites which allow you to run pings from their location. (To find them use a search engine to look for 'ping websites'.)
- By triangulating from different points it possible to get a sense of the relationship between latency and physical distance, assuming that the (approximate) geographic location of the origins and target are known. It is possible to get the physical distance (measured as the great circle distance) between the cities using a websites distance calculator (E.g. John A. Byers's site *www.wcrl.ars.usda.gov/cec/java/lat-long.htm*).
- Knowing the latency and physical distance also means you can approximate the speed of data transmission. Indeed, the use of ping has been taken further in a classroom physics experiment to calculate the speed of light! For technical details see, Lepak J. and Crescimanno M. (2002). "Speed of light measurement using ping", Report-no: YSU-CPIP/102-02, (http://arxiv.org/abs/physics/0201053).

4. Routes through the Internet

Q. *What route does my data take through the Internet?*

A. You can answer this with *traceroute*, a network engineer's tool that allows you to 'lift the lid' on the Internet and get a packets'-eye view of its working structure. It is undoubtedly the most useful tool available for Net:Geography fieldwork.

- Traceroute works in much the same way as ping but it provides much greater detail. It maps out the path that data packets take between two points on the Internet, showing all of the intermediate nodes traversed, along with an indication of the speed of travel for each segment of the journey.
- Traceroute was invented in 1988 by Van Jacobson at Lawrence Berkeley National Laboratory in the U.S. The utility often comes as part of the operating system. The Windows version is called *tracert* and is used simply by typing, at the command prompt, *tracert [internet address, e.g. jasonnolan.net]*.
- Although, traceroute is primarily for network engineers to 'debug' routing problems, it can also be used in a more tactical fashion by researchers to expose the political-economic structures of the Internet. It reveals the hidden complexity of data flows, showing how many nodes are involved (often more than twenty), the seamless crossing of oceans and national borders and the sometimes convoluted transfers through separate networks owned and operated by competing companies.
- To illustrate how traceroute 'maps' the Internet, it was used to chart the path from a PC in London to the *jasonnolan.net* website. The program works away for a few seconds as it dynamically explores the route, finding out about one node at a time. Figure 11 shows the result.
- It is important to note that the view of data routing that is 'seen' by traceroute is quite a generalised, 'high-level' summary of the network in terms of topological connections. Below the traceroute view, there is a much more detailed level of routing in terms of the geography that lies between each node, based on the types of wires and their physical pathways in the ground. Sadly, this level of detail is not measurable using Net:Geography type fieldwork techniques on the Internet itself.



```
C:\>tracert www.jasonnolan.net

Tracing route to jasonnolan.net [64.246.60.38]
over a maximum of 30 hops:

  1   191 ms   180 ms   160 ms  webport-cl2-hg9.ealing.mdip.bt.net [212.140.88.200]
  2   211 ms   180 ms   170 ms  192.168.1.38
  3   200 ms   160 ms   180 ms  interconnect5-10.ealing.fixed.bt.net [195.99.125.166]
  4   240 ms   531 ms   180 ms  core1-pos4-3.ealing.ukcore.bt.net [194.72.9.241]
  5   481 ms  1051 ms   301 ms  core1-pos10-0.redbus.ukcore.bt.net [194.74.65.254]
  6  1352 ms     *        *     lndnuk1icx1.wcg.net [195.66.224.105]
  7     *        *     2794 ms  nycmny2wcx2-pos3-1.wcg.net [64.200.87.153]
  8   901 ms   631 ms   240 ms  hrndva1wcx2-pos5-0.wcg.net [64.200.210.97]
  9   571 ms   251 ms   240 ms  hrndva1wcx3-pos9-0.wcg.net [64.200.95.74]
 10   290 ms   271 ms   260 ms  drvlga1wcx2-pos4-0.wcg.net [64.200.232.125]
 11   301 ms   250 ms   261 ms  drvlga1wcx1-oc48.wcg.net [64.200.127.29]
 12   300 ms   290 ms   291 ms  dllstx1wcx3-pos6-0.wcg.net [64.200.240.21]
 13   330 ms   271 ms   280 ms  dllstx1wcx2-pos10-0-oc48.wcg.net [64.200.110.77]
 14   331 ms   270 ms   280 ms  hstntx1wce2-pos4-0.wcg.net [64.200.240.74]
 15   331 ms   280 ms   271 ms  hstntx1wce2-everyonesinternet-gige.wcg.net [65.77.93.54]
 16     *      721 ms  2043 ms  207.218.245.113
 17   320 ms   301 ms   290 ms  jessica.cpanelserver.co.uk [64.246.60.38]

Trace complete.

C:\>_
```

Figure 11: Traceroute from London to www.jasonnolan.net.

Q. *This does not look much like a map, can you explain what it means?*

A. Yes, the end result of a traceroute does look rather cryptic at first sight, but it is in fact a kind of one-dimensional map of how the data flows, with each node traversed listed on a separate line.

- The ‘map’ gives a complete linear route listing showing how data packets travelled through the Internet starting in London and ending at Houston in the U.S.—the apparent location of the Web server which publishes *jasonnolan.net*. The three time measurements in milliseconds—such as *211ms 180ms 170ms* – are round trip times for that segment and give a useful indication of the speed of each link.
- Each node traversed is identified by its domain name and numeric IP address. Not all nodes have a domain name (e.g. *192.168.1.38*). Also, notice that many nodes have strange, long domain names (e.g. *dllstx1wcx3-pos6-0.wcg.net*). These are routing computers at the core of the Internet and their domain names are not normally seen by users, instead they are designed specifically to inform engineers running the network. With a little bit of ‘decoding’ these router domain names can yield useful information, such as the type of node hardware, the bandwidth of the link, the name of ISP that owns a node and often a node’s approximate location (usually at the city level). Fortunately, for traceroute explorers, many of the large ISPs apply a consistent naming

conventions throughout their network infrastructures (as you can see from the domain names of nodes owned by *wcg.net* in Figure 11).

- The geographic location of the node is often represented in these types of domain names as an abbreviated city name. For example, *dllstx* at the start of segments 12 and 13 (Figure 11) could sensibly be guessed to mean Dallas, Texas. Some ISPs use the familiar three letter airport identification codes (e.g. LHR for London Heathrow) as their city naming convention. (There are lists of these airport codes available on the Web, for example at, *www.orbitz.com/App/global/airportCodes.jsp*.)

Q. *How do I interpret the actual route to jasonnolan.net from the traceroute output?*

A. The first thing to note is that data travelling from London to *jasonnolan.net* had to pass through 16 intermediate network routers, to reach the end of the destination (node 17). At least three different networks were traversed - British Telecom (BT), Williams Communication Group (WCG), and Everyones Internet.

- Reading the route line by line, it begins with the first node that a user's PC is connected to the Internet, via a dial-up connection. From the domain name we can see that it belongs to *bt.net.* From local knowledge, it is know that *'ealing'* in the domain name is also an area in West London, so we can take this is an indicator of its likely geographic location.
- The next 'hop' in the journey to node 2 is rather mysterious with no domain name to decode. We have to assume it is a node within BT's network in London.
- Node 3's domain name indicates it is a another BT node in Ealing, London.
- Node 4 again says *'ealing'* and BT. The node also say *'ukcore'* which we might reasonably take to mean this node is within BT's core network for the United Kingdom.
- Node 5 is also in BT's '*ukcore*' network. Notice, the increase in latency as measured by the RTT at this point in the journey.
- At 'hop' 6 in the journey the data leaves BT's network and is handed off to another ISP called wcg.net (Williams Communication Group, now part of Wiltel corporation). The cryptic abbreviation at the start of the domain name (*'lndnuk1icx1'*) can reasonably be decoded as London, UK. The convention on this ISP's network is to start the domain name with a 4 letter abbreviation of the city, followed by a 2 letter code for country / U.S. state. Note, the big jump in RTT and the appearance of * for

two of the times (this means timed-out, no response) at this point, probably due to traffic ‘congestion’.

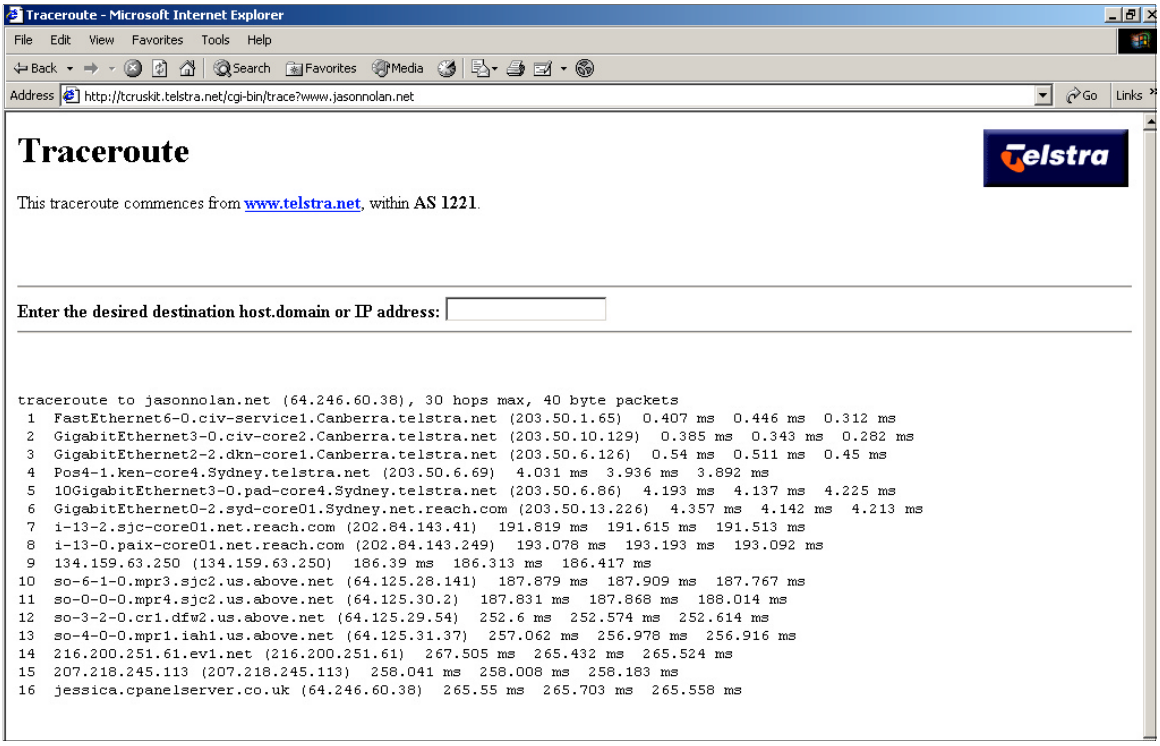
- The next segment in the journey sees the data packets cross the Atlantic to New York, most likely on an undersea fibreoptic cable. The start of the domain name for node 7 is *‘nycmny’* which can be decoded as New York City, New York. The RTT increases greatly at this point, again with two * timeoutes.
- From New York the data travels on wcg.net network to *‘hrndva’* at nodes 8 and 9, which is Herndon, Virginia (one of Washington D.C.’s satellite towns which has a lot of Internet infrastructure related companies).
- The next two steps in the journey on wcg.net’s network are in *‘drvlga’* which is somewhere in the state of Georgia. However, it is not immediately obvious which town *‘drvl’* refers to. Perhaps it is a suburb of Atlanta, the main Internet hub point for the state.
- We are now approaching our goal, as the data moves on into the state of Texas, going through Dallas (*‘dllstx’*) in nodes 12 and 13 and then to final destination, the city of Houston, Texas (*‘hstntx’*) at node 14.
- At node 15, the wcg.net network exchanges the data to a new company, everyonesinternet (Everyones Internet, Inc).
- Nodes 16 is most likely on EveryonesInternet network but does not have a domain name so it is hard to know for sure.
- Node 17, somewhat confusingly called *‘jessica.cpanelserver.co.uk’*, is the domain name of the Web server that hosts *jasonnolan.net* website. It is quite unclear why this server in Houston has a *co.uk* domain name!
- This might seem like a complicated journey. On one level it is quite a feat of routing and co-operation, but it is all in a day’s work for the Internet. These kinds of journeys happen, unseen, for the many millions of Internet users and they never need not worry about where their data flows.
- It is important to realise that Internet routing is dynamic, it can change minute by minute. The ‘map’ that is produced by traceroute is a live scan and always represents a one-off snapshot of Net:Geography at the point in time it was charted. Running the same trace at a future time is quite likely to give a different map.

Q. *Can I run traceroutes from different places?*

A. Yes, just like ping you can 'triangulate' the Internet using Web-based traceroutes.

- These make it possible run traces from many different starting points, including on different networks and in completely different continents. Web traceroute gateways are very useful for active exploration of the Internet's topology from across the globe and illustrate the degree to which routes vary.
- There are several hundred freely available Web traceroute gateways in many places. Thomas Kernen maintains a good list of them at *www.traceroute.org*.
- As an example of traceroute triangulation we ran a trace from Australia to *jasonnolan.net* using a gateway provided publicly by Telstra, the main Australian telecoms carrier. Figure 12 shows the output 'map' from the trace. (Note that the formatting of this output is a slightly different to that produced by Windows tracert.)
- Again, with a little bit of decoding work, reading line by line, it is possible to follow the data packets on this new journey. The traceroute utility is installed on a *telstra.net* server located in Canberra, Australia indicated by the domain name for node 1. The next two nodes in the trace were also within Canberra, according to their domain names.
- At node 4 the data moved a couple of hundred miles on the *telstra.net* network from Canberra to Sydney. The data is then passed through three nodes in Syndey before leaving the Telstra network for the *reach.com* network at node 6.
- The big trans-Pacific hop in the journey occurred at node seven, as the data went to *'sjc'*, the airport code for San Jose, California. Note, the marked jump in the RTTs at this point in the journey, caused in large part by the 7,500-mile distance across the Pacific Ocean.
- Node 8 on the *reach.com* network is located at *'paix'*, the name of a major Internet exchange point located in Palo Alto, California. Node 9 is cryptic. At node 10 the data left PAIX for a new network, that of *above.net*.
- Nodes 10 and 11 on Above.net's network were located in San Jose, California as indicated by the *'sjc'* codes in their domain names.
- The data moved on from California into Texas, going to Dallas-Fort Worth (*'dfw'*) at node 12. It moved onto Houston, Texas (*'iah'*) at node 13.

- The final stretch into *jasonnolan.net* took place at nodes 14 and 15, which are likely to still be in the Houston area.



Traceroute - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address http://tcruskit.telstra.net/cgi-bin/trace?www.jasonnolan.net
Go Links
Telstra

Traceroute

This traceroute commences from www.telstra.net, within **AS 1221**.

Enter the desired destination host.domain or IP address:

```
traceroute to jasonnolan.net (64.246.60.38), 30 hops max, 40 byte packets
 1  FastEthernet6-0.civ-service1.Canberra.telstra.net (203.50.1.65)  0.407 ms  0.446 ms  0.312 ms
 2  GigabitEthernet3-0.civ-core2.Canberra.telstra.net (203.50.10.129)  0.385 ms  0.343 ms  0.282 ms
 3  GigabitEthernet2-2.dkn-core1.Canberra.telstra.net (203.50.6.126)  0.54 ms  0.511 ms  0.45 ms
 4  Pos4-1.ken-core4.Sydney.telstra.net (203.50.6.69)  4.031 ms  3.936 ms  3.892 ms
 5  10GigabitEthernet3-0.pad-core4.Sydney.telstra.net (203.50.6.86)  4.193 ms  4.137 ms  4.225 ms
 6  GigabitEthernet0-2.syd-core01.Sydney.net.reach.com (203.50.13.226)  4.357 ms  4.142 ms  4.213 ms
 7  i-13-2.sjc-core01.net.reach.com (202.84.143.41)  191.819 ms  191.615 ms  191.513 ms
 8  i-13-0.paix-core01.net.reach.com (202.84.143.249)  193.078 ms  193.193 ms  193.092 ms
 9  134.159.63.250 (134.159.63.250)  186.39 ms  186.313 ms  186.417 ms
10  so-6-1-0.mpr3.sjc2.us.above.net (64.125.28.141)  187.879 ms  187.909 ms  187.767 ms
11  so-0-0-0.mpr4.sjc2.us.above.net (64.125.30.2)  187.831 ms  187.868 ms  188.014 ms
12  so-3-2-0.cr1.dfw2.us.above.net (64.125.29.54)  252.6 ms  252.574 ms  252.614 ms
13  so-4-0-0.mpr1.iah1.us.above.net (64.125.31.37)  257.062 ms  256.978 ms  256.916 ms
14  216.200.251.61.ev1.net (216.200.251.61)  267.505 ms  265.432 ms  265.524 ms
15  207.218.245.113 (207.218.245.113)  258.041 ms  258.008 ms  258.183 ms
16  jessica.cpanelserver.co.uk (64.246.60.38)  265.55 ms  265.703 ms  265.558 ms
```

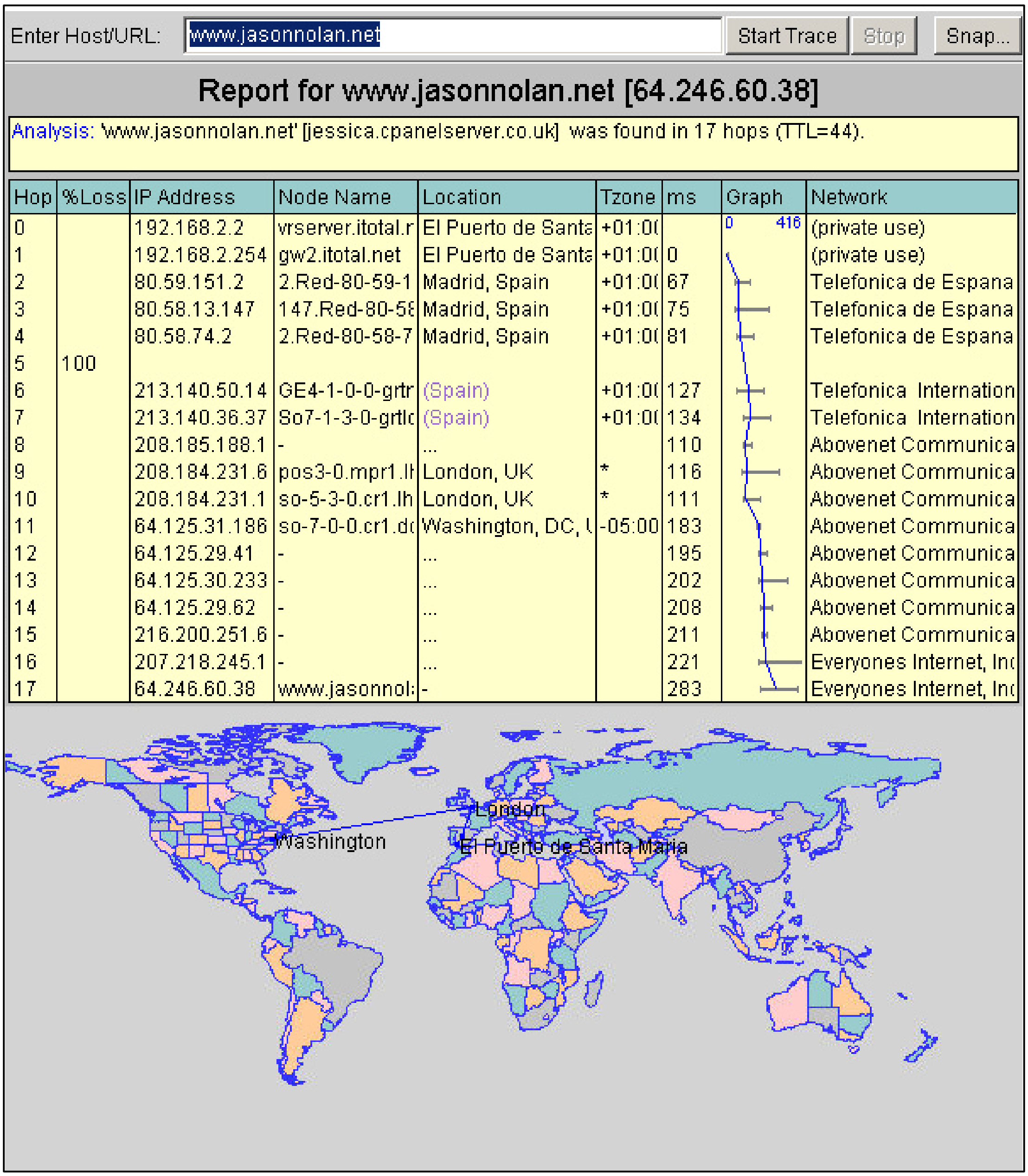
Figure 12: An example of a Web-based traceroute from Canberra, Australia to jasonnolan.net (www.telstra.net/cgi-bin/trace).

Q. *Can I geographically map traceroute output?*

A. Yes, an obvious refinement of the regular traceroute list output is to try to plot the route visually on a geographic map. There have been a number of attempts at a geographical traceroute, with varying degrees of success.

- Figure 13 is a screenshot of the best geographic traceroute program currently available, called VisualRoute (www.visualroute.com), tracing the data route from a server is Spain to *jasonnolan.net.* The top half of the display is table presentation of the trace results. The approximate locations (where known) of the routers are plotted on a rather crude map below.
- The particular advantage of this application is the ease of geographic interpretation of routing. For example, it can provide direct visual evidence of the Internet's business 'logic' of data routing following the cheapest paths rather than the geographical shortest. Much international Internet traffic is still routed through the U.S. as the cheapest means of transit between regions. This can result in sometimes quite anomalous looking, circuitous routes being chosen.

- However, automated geographic traceroute is far from perfect. It is very hard for software to reliably ‘decode’ the router domain names as this often requires local knowledge, human intuition and a bit of guesswork. The large number of gaps in the ‘Location’ column of the traceroute results table in Figure 13 clearly show the limitations.



Report for www.jasonnolan.net [64.246.60.38]

Analysis: 'www.jasonnolan.net' [jessica.cpanelserver.co.uk] was found in 17 hops (TTL=44).

Hop	%Loss	IP Address	Node Name	Location	Tzone	ms	Graph	Network
0		192.168.2.2	vrserver.itotal.r	El Puerto de Santa	+01:0(		0 416	(private use)
1		192.168.2.254	gw2.itotal.net	El Puerto de Santa	+01:0(	0		(private use)
2		80.59.151.2	2.Red-80-59-1	Madrid, Spain	+01:0(	67		Telefonica de Espana
3		80.58.13.147	147.Red-80-58	Madrid, Spain	+01:0(	75		Telefonica de Espana
4		80.58.74.2	2.Red-80-58-7	Madrid, Spain	+01:0(	81		Telefonica de Espana
5	100							
6		213.140.50.14	GE4-1-0-0-grtr	(Spain)	+01:0(	127		Telefonica Internation
7		213.140.36.37	So7-1-3-0-grtlc	(Spain)	+01:0(	134		Telefonica Internation
8		208.185.188.1	-	...		110		Abovenet Communica
9		208.184.231.6	pos3-0.mpr1.lh	London, UK	*	116		Abovenet Communica
10		208.184.231.1	so-5-3-0.cr1.lh	London, UK	*	111		Abovenet Communica
11		64.125.31.186	so-7-0-0.cr1.dc	Washington, DC, (	-05:00	183		Abovenet Communica
12		64.125.29.41	-	...		195		Abovenet Communica
13		64.125.30.233	-	...		202		Abovenet Communica
14		64.125.29.62	-	...		208		Abovenet Communica
15		216.200.251.6	-	...		211		Abovenet Communica
16		207.218.245.1	-	...		221		Everyones Internet, Inc
17		64.246.60.38	www.jasonnol:	-		283		Everyones Internet, Inc

Figure 13: An example of a geographic traceroute produced using VisualRoute.

Q. *What else can I tell from traceroute results?*

A. There are practical things you can use traceroute data for. It can also used for more political 'debugging' of the Internet's structure.

- Traceroute can be really useful for deducing the approximate location of Internet hosts (such as websites) in terms of 'hardware geography'. The output can tell you the location and identity/owner of the 'upstream' network provider even if the final destination of the server is unclear. If the data travels into a certain city and does not leave it again, it is probable that the target is located there. For example, deducing that *jasonnolan.net* is published from a Web server in Houston, Texas. Also, the 'upstream' network providers may keep logs for identifying the host that is of interest (this is of particular concern for law enforcement agencies in tracing the source of illegal activities).
- Traceroute has also been used in physics classroom experiment to measure the size of the earth. See the paper, Kicovic S., Webb L., and Crescimanno M. (2002). "Measuring the earth with traceroute" (http://arxiv.org/abs/physics/0208087).
- Running multiple traceroutes to lots of different points across the Internet has also been used to gather data to chart the topology of the core of Internet. The results of which have been visualised as huge abstract graphs, providing some of the most evocative representations of Net:Geography seen. See for example Lumeta's Internet Mapping Project (http://lumeta.com/mapping.html); their technical paper gives for further details: Branigan, S., Burch, H., Cheswick, B., and Wojcik, F. (2001). "What Can You Do with Traceroute?". *Internet Computing*, September/October 2001, Vol. 5, No. 5, page 96 (http://computer.org/internet/v5n5/index.htm).
- Traceroute can also reveal something of the hidden political economy of Internet. The patterns of traffic routing shows transit agreement and mutual peering relationship between competing companies. Details on these deal are often deemed commercially confidential but are revealed by necessity in how and where the actual networks interconnect to share data. The routing of traffic reveals the structuring of business relationships in terms of who connects to who and the hierarchy of these connections (from local to centre to local again) shows the power geometries of the information economy. It can also show which telecommunications carriers dominate the transfer of traffic between countries and between continents. These companies are likely to be

influential in the structuring of global communications and tracerouting could provide an alternative way to quantify the extent of power.

- Traceroute can show potential vulnerabilities in the structure of the Internet. Are there particular choke points in the routing of data flows? Is there only single route into a region or between two cities?
- Lastly, the output from traceroute provides a useful way to assess the number of international borders crossed and determine which different territories (i.e. separate legal jurisdictions) the data transits. The more 'points of contact' in the flow from origin to target, the more potential there is that Internet traffic could be intercepted and subjected to local regimes of monitoring, filtering, censorship and data retention. For example, does your email message transit through a third-party nation that has hostile intentions. Particularly in regions of conflict, being able to identifying territories that are transited might be vitally important in terms of the security of communication. Does an email to someone in Palestine transit through Israel?

Conclusion

Q. *So what is the future of Net:Geography fieldwork?*

A. As the Internet grows in size, expands in scope and becomes increasingly embedded as a banal and invisible background to everyday living it becomes more important to understand its politics. We would argue that understanding the geographies of the Internet, through Net:Geography fieldwork using the techniques and tools described here, provides one of the most valuable avenues into network politics, allowing you to gather information and critically questions network operations first hand.

- Net:Geography fieldwork is likely to become easier as new and more powerful software tools for scanning the structure of Internet become available. This will be a benevolent outcome of the experience in the design and the countering the current plague of Internet worms and viruses. As search engine tools develop they will increasingly provide revealing new ways to do Net:Geography fieldwork in analysing the information structures of the Internet. Of course, researchers will continue to have to tread carefully the ethical boundaries between critical fieldwork and potentially criminal hacking.
- Yet, at the same time, Net:Geography fieldwork is going to be harder and riskier to do. Individual networks on the Internet are increasing being designed and operated in a

much more closed fashion. For example, the university networks of the authors have recently blocked ping and traceroutes as a security precaution against malicious intrusion. Other areas of the Internet are also using the cover of greater security as a way to develop more proprietary and profitable business operations. While many users of the Internet, for example on peer-to-peer networks, are likely to be using in future software tools that encrypt and mask their activities and their locations to preserve confidentiality of communication in the face of invasive monitoring by corporations and governments. This will also have a side effect of frustrating Net:Geography fieldwork. Lastly, there is a definite ‘chilling effect’ on anyone who without official authorisation showing ‘unhealthy’ interest in details on the location of infrastructures and technicality of network operations. So don’t be surprised if doing some Net:Geography fieldwork makes you a subject of suspicion.