VISUALIZATION FOR IT - SECURITY

L. Aaron Kaplan (kaplan@cert.at)

http://CERT.AT
OVERVIEW

• Motivation
• Target Group
• 5 Minutes of design background for techies
• Tools
• DNSviz and Flows
Why am I holding this talk?
Why am I holding this talk?
Why am I holding this talk?
Why am I holding this talk?
Why am I holding this talk?
Why am I holding this talk?
Overview

• What will you get out of it?
  • Quick IT–security visualization skills with 5 tools
  • Understanding the basic visualization cycle
  • Initial good results in < 1 day
  • Really good results in 10+ years ;-(
CERT.at, Austria

CERT.at is part of NIC.at, the Austrian domain registry. CERT.at is the national CERT.

Austria is in Europe, but we definitely like the friends from AUSCert and down under.

Vienna, Austria is where we will have our next FIRST conference 2011.

German is spoken in Austria.

Our neighbouring countries are: Hungary, Slovenia, Germany, Switzerland, Slovakia, Czech Republic, Italy, Liechtenstein.
CERT.at, Austria

• CERT.at is part of NIC.at, the Austrian domain registry. CERT.at is the national CERT
• Austria is in Europe, but we definitely like the friends from AUSCert and down under
• Vienna, Austria is where we will have our next FIRST conference 2011
• German is spoken in Austria
• Our neighbouring countries are: Hungary, Slovenia, Germany, Switzerland, Slovakia, Czech Republic, Italy, Liechtenstein
OVERVIEW

- Motivation
- Target Group
- 5 Minutes of design background for techies
- Tools
- DNSviz and Flows
Motivation
Motivation

- “A picture is worth 1000 log records” (R. Marty)
• “A picture is worth 1000 log records” (R. Marty)
• We have too much data, info explosion
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- High broadband path to your brain
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/father/mother/partner...
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/father/mother/partner...
- ... and helps them understand that you need to save the internet first
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/ father/mother/partner...
- ... and helps them understand that you need to save the internet first
- gives new insights -> explore data
Motivation

• “A picture is worth 1000 log records” (R. Marty)
• We have too much data, info explosion
• **High broadband path to your brain**
• People “get it”
• Visualization can explain it all to your grandpa/father/mother/partner...
• ... and helps them understand that you need to save the internet first
• gives new insights –> explore data
• gives us an overview
Motivation

- “A picture is worth 1000 log records” (R. Marty)
- We have too much data, info explosion
- **High broadband path to your brain**
- People “get it”
- Visualization can explain it all to your grandpa/father/mother/partner...
- ... and helps them understand that you need to save the internet first
- gives new insights -> explore data
- gives us an overview
- sells your services
OVERVIEW

• Motivation
• Target Group
• 5 Minutes of design background for techies
• Tools
• DNSviz and Flows
Target groups

• Users
• Management, Sales, Politicians

• Operational staff
• Researchers
Target groups

- Users
- Management, Sales, Politicians
- Operational staff
- Researchers

source: CAIDA.org
Target groups

- Users
- Management, Sales, Politicians
- Operational staff
- Researchers

Conficker Eye Chart
OVERVIEW

- Motivation
- Target Group
  - 5 Minutes of design background for techies
- Tools
- DNSviz and Flows
Some design background

• One of the leading persons in the field right now: Edward Tufte

• Learned a lot from Otto Neurath: “Isotypes” in Vienna in the early 1900s

• First invention of “icons”. Idea: educate the illiterate working class population in Europe w.r.t basic economics relationships
Otto Neurath’s Isotype

Kriegsverluste

1525 Bauernkrieg

1812 Zug Napoleons nach Russland

1861-65 Amerikanischer Sezessionskrieg

1870-71 Deutsch-Französischer Krieg

1914-18 Weltkrieg

Jede einzelne Figur 50.000 Mann Überlebende
Jedes einzelne Kreuz 50.000 Tote

Deutsche
Franzosen

Entente
Zentralmächte

Die Figuren sind auf 300.000 Mann abgerundet, die Verluste auf 3 Prozent.
Otto Neurath’s Isotype

Wohndichte in Großstädten

Einige Weltstädte

New York
London
Paris
Berlin
Chicago
Wien

Die deutschen Großstädte über 400 000 Einwohner

Berlin
Hamburg
Köln
München
Leipzig
Dresden
Breslau
Frankfurt a. M.
Essen
Düsseldorf
Hannover

Anordnung der Städte nach ihrer Größe. Anfang 1929
Maschinenausfuhr vor dem Krieg und jetzt

Durchschnitt
1909 - 1913

1928

Jeder Kreis 100 Millionen Mark

Die Längen der Rundbildern geben einen ungefähren Vergleich der Mengen der ausgeführten Maschinen. Die Kupferkosten des Goldes ist auf ca. 20% gesunken.
Otto Neurath's Isotype

Entwicklung der Eisenbahnen

1825
1851
1881
1901
1913
1926

Übriges Amerika
U. S. A. Europa
U.d.S.S.R. Übrige Welt

50,000 Streckenkilometer

Angelernt für den Bibliographischen Institut AG, Leipzig
Geschichte- und Wirtschaftssysteme in Welt.
Handelsmarinen der Erde

1850

1900

1913

1929

Jedes Schiff 5 Millionen Bruttoregistertonnen
Otto Neurath’s Isotype

Waldbestand in Eurasien
Modern day examples

How many people are connected to the internet?

In 2009, we had approximately 6,767,805,208 people on the earth from those, 1,802,330,457 have internet access which makes it 26.6% or one quarter of the world population.

(source: http://www.internetworldstats.com/stats.htm)
Modern day examples of waste of resources by spam and a spammer's income

Handling, sending, receiving and filtering out spam takes up the power of 2.4 million US houses or 1 nuclear power plant.

Each mail represents 1 trillion spam mails
Each house represents 100,000 US houses
An average nuclear power plant produces 2,500 Mwatts

Good Emails 2%
Spam 98%
Orders from Spam 0.02%

US household per month
Spammer per day

(source: McAffee CO2 Impact of Spam + NY Times)
Making users understand IT security

Making users understand IT security

sources: Arbor ATLAS, spamcop.net
Making users understand IT security

sources: Arbor ATLAS, spamcop.net
OVERVIEW

• Motivation
• Target Group
• 5 Minutes of design background for techies

• Tools
• DNSviz and Flows
• Graphviz
• Maxmind GeoIP
• Logster
• Unix wizardry
• Google Earth
• Gapminder
• Processing.org
• Outlook: Davix
• based on research @ AT&T Labs
• Syntax:

```
digraph {
    A -> B;
    A -> C [label="foo"];}
```

dot -T png -o out.png \inputfile.dot
Maxmind GeoIP

- [http://maxmind.com](http://maxmind.com)
- cityLite DB is usually enough

```perl
my $gi = Geo::IP->open("/home/aaron/GeoLiteCity.dat", GEOIP_STANDARD);

# ----------------- functions ----------------
# input : ip
# output: array [countrycode, city, lat, lon]
sub ip2geolocate {
    my $ip = $_[0];
    my @ret;
    my $record = $gi->record_by_name("$ip");
    @ret = ( $record->latitude , $record->longitude) ;
    return(@ret);
}
```
Tools: Logster

- Logster by Clarified Networks
- Input format: Apache log file format
- Output: movie. Can screen capture
Tools: Logster

Input format: Apache log file format

Output: Movie. Can screen capture
Tools: Gapminder
Tools: Google Earth

- format: KML. Well documented.
- Head section
- Placemarks
Tools: Unix filters

- Use Unix tools to quickly get a grasp of the trends
  - `cut -d ";" -f 5 | sort | uniq -c | sort -rn`
  - `gnuplot`

  `plot "myfile.csv" using 1 with boxes`
Albert-László Barabási made them famous.

Some property is distributed by an inverse power law formula: 
\[ P(k) \sim \frac{1}{k^\gamma} \quad (2 < \gamma < 3) \]

“fractal”

“internet–ish”

“biological”

“not again–ish”
TOOLS: Processing.org

- Invented by Ben Fry, Casey Reas @MIT
- Basic idea: easy IDE for Java 3D/OpenGL programing. Lots of examples, openprocessing.org
- Includes a rich API:
  - sockets
  - DB connections
  - serial I/O
  - sound, etc.

```java
import processing.opengl;

boolean blueAlgorithm = false;

int border=10;
int num_circles=-1;

textSize = 16;

float mappCircleArea = 4000.0;
float mappCircleRadius;
float mappDistance;

float angularStep = 2*PI/360.0*360; // 3 degrees
float defaultTextAngle = PI/3; // how to rotate the text by default

// arrays for the circles
```
Processing example: circular layout
Other processing Examples

- Esfera
- Registrymon
Outlook: DAVIX

- ISO image on [http://www.secviz.org](http://www.secviz.org)
OVERVIEW

• Motivation
• Target Group
• 5 Minutes of design background for techies
• Tools
  • DNSviz and Flows
DNS
DNS
DNS for IT security viz

- Authoritative
  - anycast utilization
  - sinkhole
  - detect misconfigurations

- Registrars
  - registrar mon
  - monitor DNS tasting

- Recursor
  - DNS as IDS
  - mapping misconfigs
  - open recursor map
  - flow mapping

- Stub
  - trace malware’s gethostbyname() calls
Idea list DNS and IT security viz

- Authoritative Nameservers:
  - you don’t see much at the authoritative NS
  - TTLs are wrong
  - other misconfigurations
  - But – idea: Spam for a newly registered domain should be a spike. But can we filter it out from the noise?
- Anycast effectiveness (c.f. CAIDA paper)
- **Sinkholing** works!
Idea list DNS and IT security viz

• Registry / Registrars:
  • from registry’s perspective: track your resellers. How “clean” is a registrar?
  • monitor DNS tasting. Find domain catchers.
• Recursors:
  • passive DNS
  • DNS “netflow” (“passive DNS++”)
  • DNS as IDS (<– Google talk today!)
  • log/visualize localhost/bogus/bogon answers!
  • fastflux
  • monitor TXT record answers
  • map (maliciously) open recursors
Idea list DNS and IT security viz

• Stub resolvers:
  • trace malware’s gethostbyname() syscalls (Minibis)
  • idea: outgoing FW + logster for the stub / PC
DNS netflow example

- Done in Processing
- data: tcpdump -ni eth0 port 53 and src = ...
- filter out local queries
- find all nameservers which are queried
- aggregate(!) + transform via perl script to...
- format:
  lat srcip; lon srcip; lat dstip; lon dstip; amount
- aggregation factor:
  - aaron@lair:~$ wc -l outgoing-without-ports.txt
  - 100000 outgoing-without-ports.txt
  - aaron@lair:~$ wc -l flows-lat-lon.txt
  - 28948 flows-lat-lon.txt
- source code demo?
DNS netflow
source: Duane Wessels, measurement factory
SIG? Data exchange?
Thanks!

annapetukhova.com, processing.org, Otto Neurath