

# Greg Conti

United States Military Academy, West Point



BLACK HAT BRIEFINGS

## Beyond Ethereal: Crafting A Tivo for Security Datastreams

Ethereal is a thing of beauty, but ultimately you are constrained to a tiny window of 30-40 packets that is insufficient when dealing with network datasets that could be on the order of millions of packets. In addition, it only displays traffic from packet captures and lacks the ability to incorporate and correlate other security related datastreams. In an attempt to break from this paradigm, we will explore conceptual, system design and implementation techniques to help you build better security analysis tools. By applying advanced information visualization and interaction techniques such as dynamic queries, interactive encoding, semantic zooming, n-gram analysis and rainfall visualization you will gain far more insight into your data, far more quickly than with today's best tools. We will discuss lessons learned from the implementation of a security PVR (a prototype will be released) and explore additional topics such as using visual techniques to navigate and semantically encode small and large binary objects, such as executable files, to improve reverse engineering. To get the most out of this talk you should have a solid understanding of the OSI model and network protocols.

*Greg Conti is an Assistant Professor of Computer Science at the United States Military Academy. He holds a Masters Degree in Computer Science from Johns Hopkins University and a Bachelor of Science in Computer Science from the United States Military Academy. His areas of expertise include network security, information visualization and information warfare. Greg has worked at a variety of military intelligence assignments specializing in Signals Intelligence. Currently he is on a Department of Defense Fellowship and is working on his PhD in Computer Science at Georgia Tech. His work can be found at [www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti) and [www.rumint.org](http://www.rumint.org).*



## Beyond Ethereal: Crafting A Tivo for Security Datastreams

Gregory Conti  
[www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti)  
[conti@cc.gatech.edu](mailto:conti@cc.gatech.edu)

## Disclaimer



The views expressed in this presentation are those of the author and do not reflect the official policy or position of the United States Military Academy, the Department of the Army, the Department of Defense or the U.S. Government.

image: <http://www.leavenworth.army.mil/usdb/standar#%20products/vtdefault.htm>

## Why?

- Way too much data from a wide variety of sources
- Complement Ethereal
  - cross cue
  - provide context / big picture
- Facilitate high level discovery - low level analysis
- Provide valuable analytic tool to analysts
- Help communicate results to wide variety of audiences
- Better observe network and intruder behavior

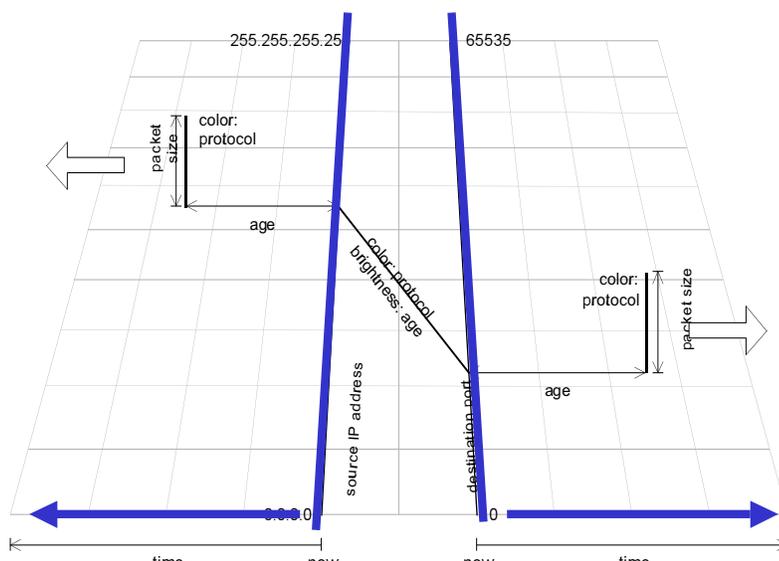
***information visualization is the use of interactive, sensory representations, typically visual, of abstract data to reinforce cognition.***

[http://en.wikipedia.org/wiki/Information\\_visualization](http://en.wikipedia.org/wiki/Information_visualization)

## Goals

- Provide world view to packet view
- Record and playback data
- Easily incorporate carefully crafted windows on the data (visualizations) to meet specific needs that aren't being addressed with current manual and machine tools.
- Incorporate all security related data sources
  - passive and active
- Scale from individual to enterprise
- Speed training
  - Dynamically create "smart book" pages with analyst markup
- Allow interactive exploration of data through such techniques and interactive encoding and filtering
  - Dynamically create filters for other tools
- Maximize customization

## Overview of secvis

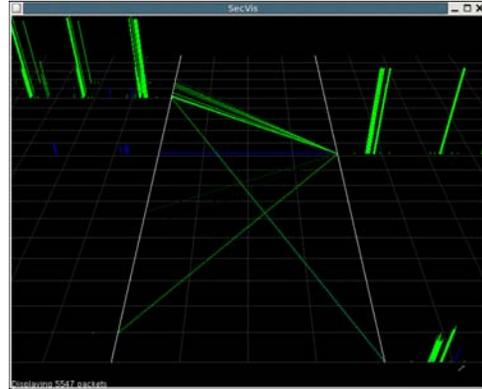


*digital self defense*

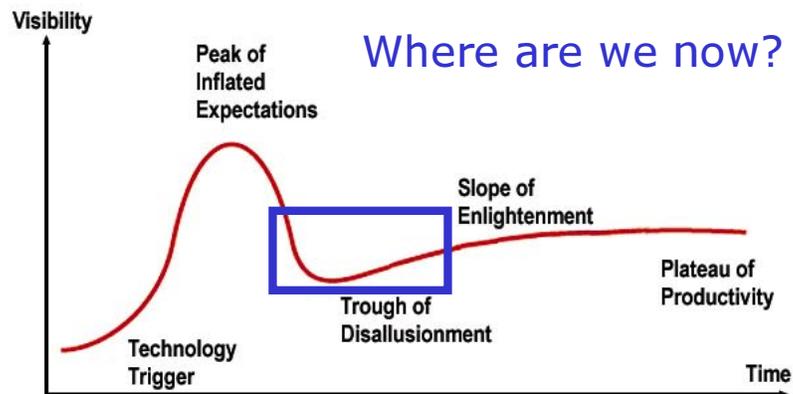
## SecVis Demo

(one possible window)

- Sven Krasser design and implementation lead
- Code on CD
  - Caveats
  - Thanks to NETI@Home
  - Released under GPL
- See the research paper for more information



## Gartner's Hype Cycle



-Gartner Group

Thanks go to Kirsten Whitely for the Gartner curve idea  
<http://java.sun.com/features/1998/03/images/year3/original/gartner.curve.jpg>

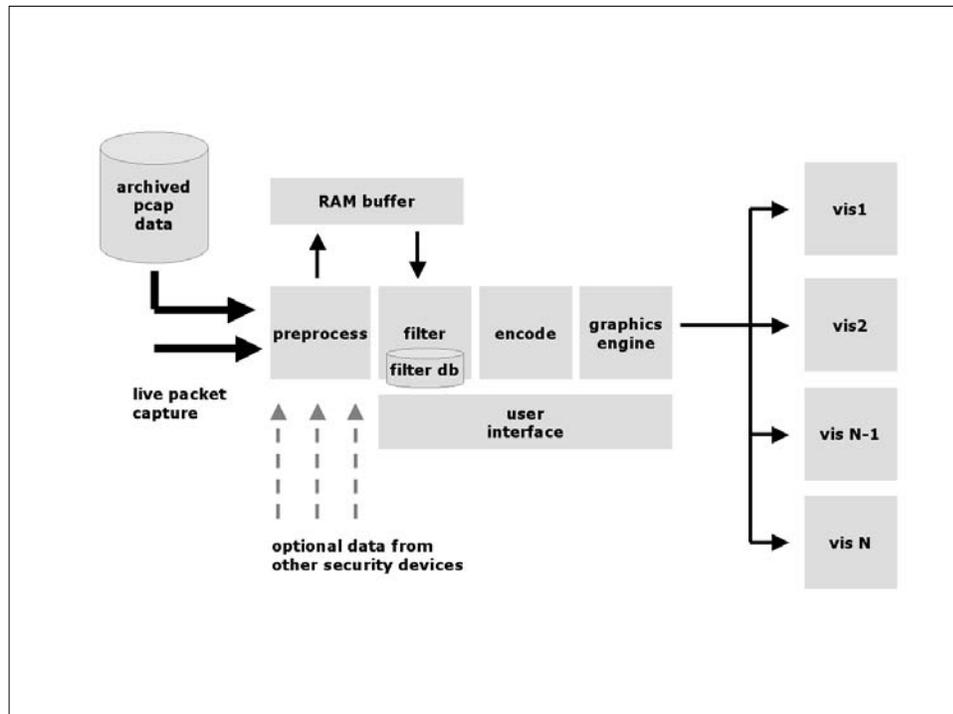
## System Characterization

- Passive vs. active data collection
  - Passive examples (firewall logs or packet capture)
  - Active examples (Internet mapping project)
- Across the spectrum from real time to offline
- Interactive exploration vs. static display
  - Granularity of interaction
  - Customizability of interaction
- Single data source vs. multiple data sources
- Information density
- Number of visualizations of data
  - Granularity of data dissection
- Applicability of techniques for given tasks
- System performance
  - Is it CPU bound, Memory bound or **Human bound**
- System security
  - Can the system and/or the dataflow be attacked

## Information Density Comparison

(graphical vs. text)

<u>Graphical</u>	<u>ASCII</u>	<u>Hex</u>
1 bit per pixel	15x	45x
8 bits per pixel	120x	360x
16 bits per pixel	240x	720x
24 bits per pixel	360x	1080x
32 bits per pixel	480x	1440x



## Potential DataStreams

- Traditional
  - pcap
  - snort
  - syslog
  - firewall logs
  - anti-virus
  - reconstruct streams
  - ...
- Less traditional
  - p0f
  - IANA data (illegal IP's)
  - reverse DNS
  - local data (unassigned local IPs)
  - inverted snort
  - active tools (e.g. nmap)
  - ...

## Data Combinations

- All parameters
- Note that all combinations are possible

- packet length (from Winpcap)
- Ethertype
- IP Transport Protocol
- Source/Destination IP
- TTL
- IP Header Len
- IP Version
- IP Diff Services
- IP Total Length
- IP Identification
- IP Flags
- IP Fragment Offset
- IP Header Checksum
- UDP Source/Destination Port
- TCP Source/Destination Port

## Methodology

- Work through slices of network traffic
- Take advantage of what the human is good at
- Create and share filters
  - toward network squelch
- Maximize customization and interaction
- Allow user to focus on what is interesting
- Knowledge discovery
- Help highlight what is interesting
- Easily drop in different windows on network traffic
- Look at traffic from different perspectives

## Design

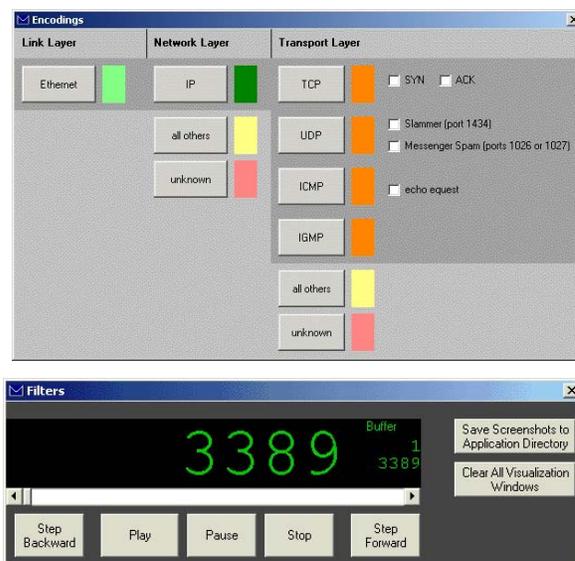
- Multiple coordinated views
- Stateless
- Buffer 100K packets at a time
- No plotting in background
- Global and visualization specific interaction
- PCAP file conversion utility required (for now)
- Visualize when appropriate
- Provide useful interactive filtering and encoding
- Apply advanced techniques

## RUMINT Main Screen



- Provide quick overview with minimal clutter
- Thumbnails act as menu
- Why "RUMINT"

## Key #1: Interaction



## Key #2: Filtering

- Internet background radiation paper
- slammer
- window sizes
- create, save and share
- flat file
- analyst comments (annotate)
- checksum errors
- TTL
- TCP flags
- band pass, inverted band pass,
- suppress repetitions

## For More Information...

- **Dynamic Queries**
  - Ben Shneiderman. <http://www.cs.umd.edu/hcil/spotfire/>
- **Requirements and Tasks**
  - Goodall. User Requirements and Design of a Visualization for Intrusion Detection Analysis
  - Komlodi, Goodall and LuttersAn Information Visualization Framework for Intrusion Detection. <http://userpages.umbc.edu/~jgood/publications/komlodi-chi04.pdf>
- **Semantic Zoom**
  - Bederson, et al., "Pad++: A Zoomable Graphical Sketchpad for Exploring Alternate Interface Physics," Journal of Visual Languages and Computing, 1996, Volume 7, pages 3-31. <http://citeseer.ist.psu.edu/bederson95pad.html>
- **Noise in Internet Data**
  - Pang, Yegneswaran, Barford, Paxson and Peterson. Characteristics of Internet Background Radiation. [www.icir.org/vern/papers/radiation-imc04.pdf](http://www.icir.org/vern/papers/radiation-imc04.pdf)
  - Grizzard, Simpson, Krasser, Owen and Riley. Flow Based Observations from NETI@home and Honeynet Data. [www.ece.gatech.edu/research/labs/nsa/papers/neti-honey.pdf](http://www.ece.gatech.edu/research/labs/nsa/papers/neti-honey.pdf)
- **Automatic Filter Generation**
  - Lakkaraju, Bearavolu, Slagell and Yurcik. Closing-the-Loop: Discovery and Search in Security Visualizations. [http://www.ncassr.org/projects/sift/papers/westpoint05\\_closing-the-loop.pdf](http://www.ncassr.org/projects/sift/papers/westpoint05_closing-the-loop.pdf)
- **Human in the Loop Systems**
  - Korzyk and Yurcik. On Integrating Human In the Loop Supervision into Critical Infrastructure Process Control Systems. [www.ncassr.org/projects/sift/papers/astc2002\\_humaninloop.pdf](http://www.ncassr.org/projects/sift/papers/astc2002_humaninloop.pdf)
  - Su and Yurcik. "A Survey and Comparison of Human Monitoring of Complex Networks." <http://www.ncassr.org/projects/sift/papers/iccrts05.pdf>

## Binary Rainfall Visualization

(single packet)

Bits on wire...

0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

View as a 1:1 relationship (1 bit per pixel)...

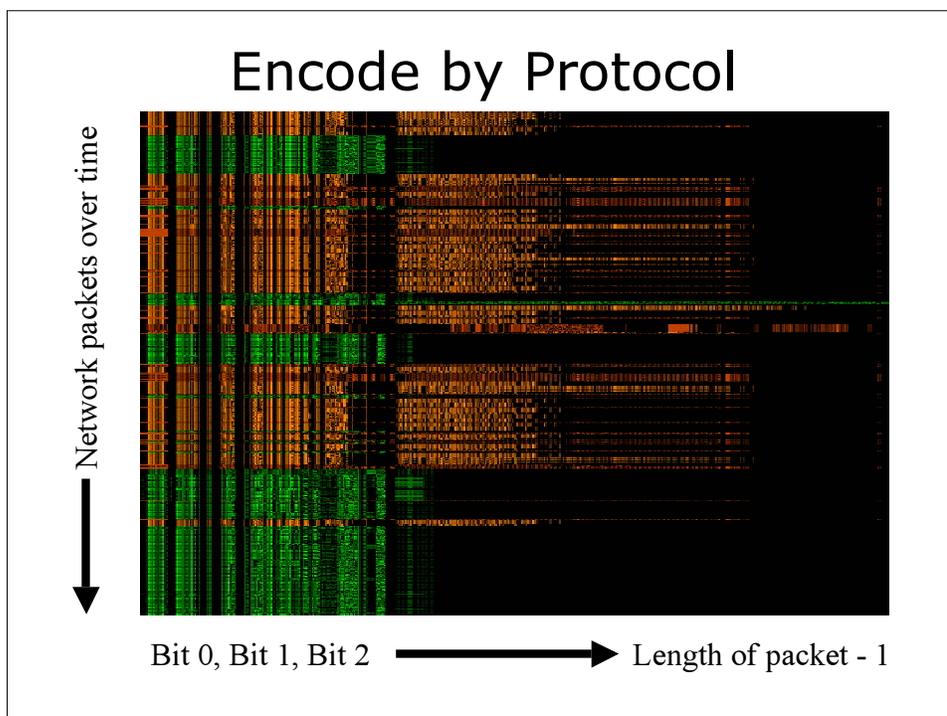
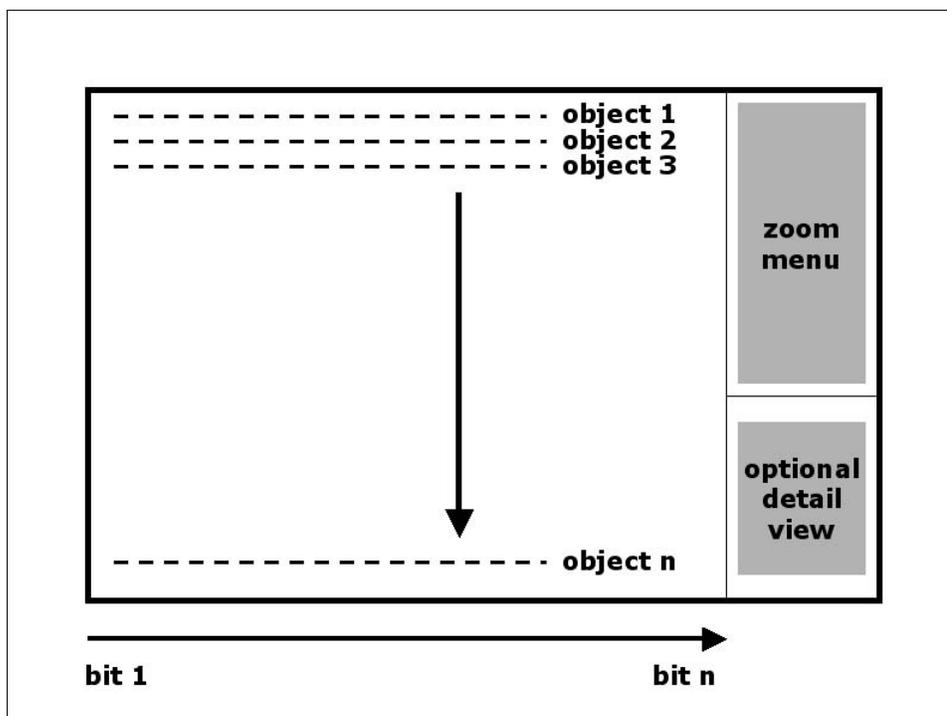
0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

View as a 8:1 relationship (1 byte per pixel)...

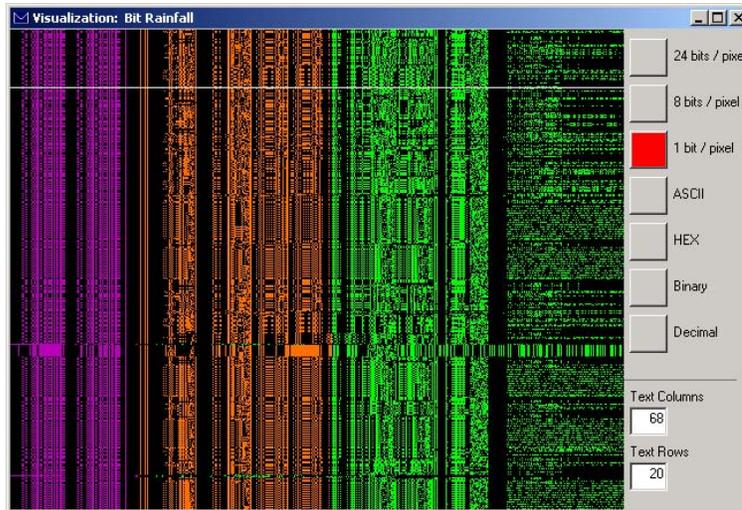
0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

View as a 24:1 relationship (3 bytes per pixel)...

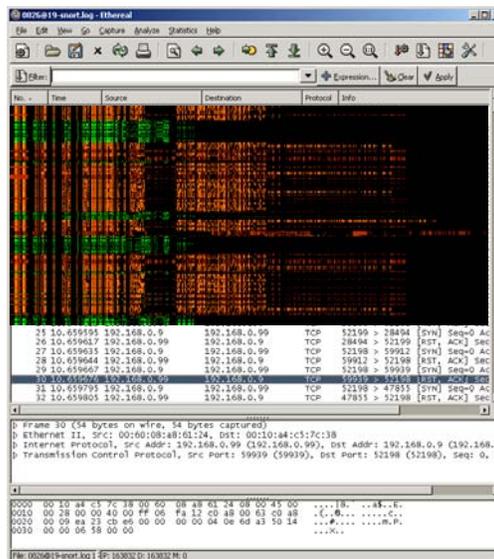
0	1	1	0	1	1	1	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



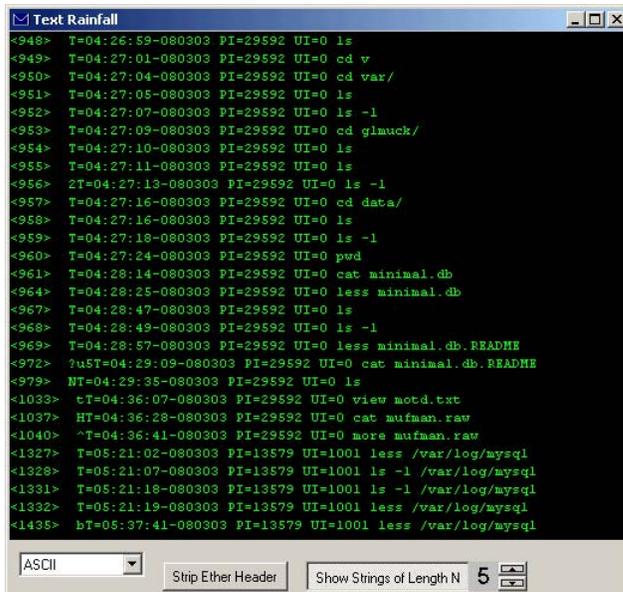
## Encoding Headers



## Navigation

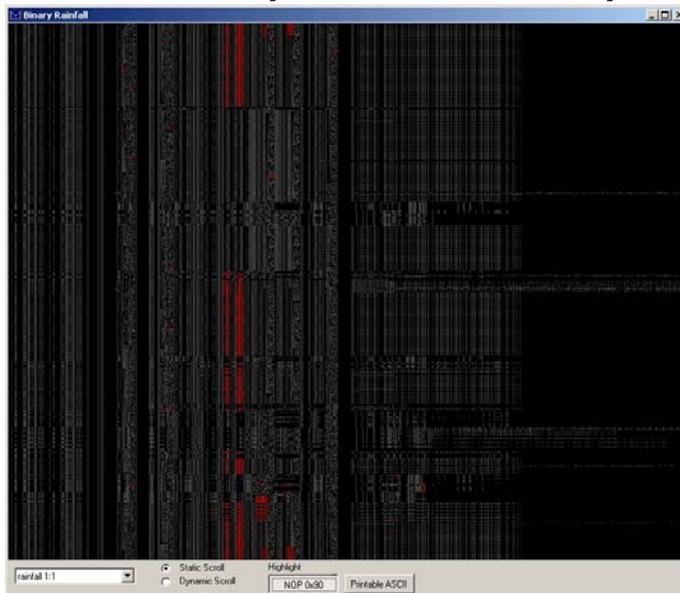


## On the fly strings



dataset: Defcon 11 CTF

## On the fly disassembly?



dataset: Honeynet Project Scan of the Month 21

## A Variant: Visual Exploration of Binary Objects

IDA Pro - [File] [Edit] [View] [Tools] [Settings] [Help]

Compare: %s %s  
Usage: %s [options] [filename],.0hh

Hex View

```
00401880 6F 6D 78 61 72 05 00 00 00 25 73 20 35 73 00 00 00 "compare.%s %s"
00401885 55 73 61 67 65 30 20 20 20 25 73 20 58 6F 7D 7A "Usage: %s [opti"
00401890 59 47 6E 73 2D 78 55 65 69 65 65 65 61 6D 65 5D "(args) [filename]"
00401895 04 04 04 47 65 6E 65 72 64 6E 20 61 78 7A 69 6F "General optio"
004018A0 6F 73 30 00 00 00 30 30 30 30 30 30 30 30 30 30 "Options:-----"
004018A5 30 30 30 30 00 00 20 20 20 20 20 20 20 20 20 20 "-----"
004018B0 20 20 20 20 20 20 20 20 20 20 5A 68 65 20 69 6E 70 " The inp"
004018B5 75 7A 20 69 73 20 3C 73 7A 6A 69 6E 3E 20 69 6E "ut is <stdin> in"
004018C0 73 7A 65 61 64 20 6F 66 20 61 20 66 69 6E 6E 6E 6E "stead of a file."
004018C5 00 00 20 68 20 7C 7A 20 3E 20 7C 20 20 20 68 65 "t -h | -? | - h?"
004018D0 6C 70 20 20 50 72 69 6E 7A 20 68 65 6C 70 2E 00 "ip Print help."
004018D5 00 20 76 20 7C 20 20 76 65 72 73 69 6F 6E 20 20 "u | -version"
004018E0 20 20 58 72 69 6E 7A 20 76 65 72 73 69 6F 6E 6E " Print version"
004018E5 20 61 6E 6A 20 63 6F 70 79 72 69 67 68 7A 2E 00 " and copyright."
004018F0 00 20 73 69 6E 65 6E 7A 20 20 20 20 20 20 20 20 " -silent"
```

<http://www.datarescue.com/ibase/>

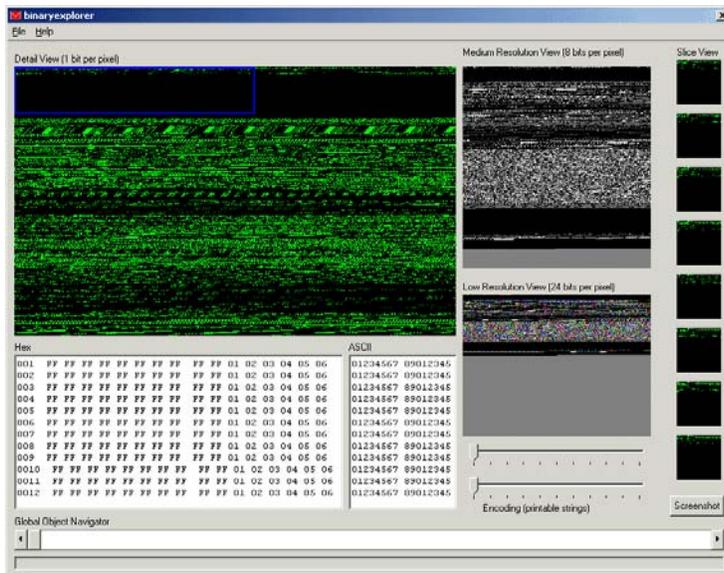
## Textual vs. Visual Exploration

binexplorer.exe

```
0000000b: 4D 5A 90 00 03 00 00 04 00 00 00 FF FF 0C
0000010b: B8 0D 00 00 00 00 00 00 40 00 00 00 00 00 00
0000020b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0000030b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0000040b: 0E 1F 8A 0E 00 84 09 CD 21 8B 01 4C CD 21 54
0000050b: 69 73 20 70 72 6F 67 72 61 69 20 63 61 6E 61
0000060b: 74 20 62 65 20 72 75 6E 20 69 6E 20 44 4F 53
0000070b: 6F 6F 64 65 2E 0D 0D 0A 14 00 00 00 00 00 00
0000080b: CD C2 79 DA 89 A3 17 89 89 A3 17 89 89 A3 17
0000090b: 0A BF 19 89 88 A3 17 89 80 8C 1E 89 8D A3 17
00000A0b: 60 BC 1A 89 8D A3 17 89 52 69 63 69 89 A3 17
00000B0b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
00000C0b: 00 00 00 00 00 00 00 00 50 45 00 00 4C 01 01
00000D0b: AD 7B 3E 41 00 00 00 00 00 00 00 00 00 00 00
00000E0b: 08 01 04 00 00 90 00 00 00 30 00 00 00 00 00
00000F0b: 2C 13 00 00 00 10 00 00 60 00 00 00 00 00 4C
0000100b: 00 10 00 00 00 10 00 00 04 00 00 00 01 00 00
0000110b: 04 00 00 00 00 00 00 00 90 00 00 00 10 00 00
0000120b: 15 08 01 00 02 00 00 00 00 00 10 00 10 00 00
0000130b: 00 00 10 00 00 10 00 00 00 00 00 00 10 00 00
0000140b: 00 00 00 00 00 00 00 00 34 5A 00 00 28 00 00
0000150b: 00 00 00 00 EC 00 00 00 00 00 00 00 00 00 00
0000160b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0000170b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0000180b: 00 00 00 00 00 00 00 00 00 00 00 00 00 00 00
0000190b: 00 00 00 00 00 00 00 00 00 38 02 00 20 00 00
```



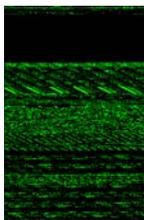
# binaryexplorer.exe



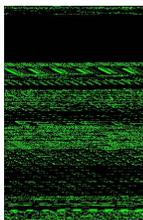
## Comparing Executable Binaries

(1 bit per pixel)

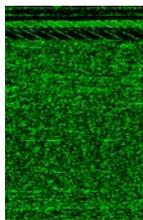
rumint.exe  
(visual studio)



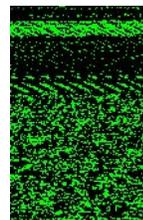
visualexplorer.exe  
(visual studio)



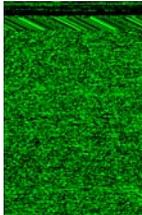
calc.exe  
(unknown compiler)



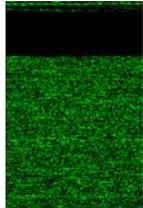
regedit.exe  
(unknown compiler)



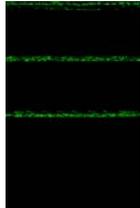
mozillafirebird.exe  
(unknown compiler)



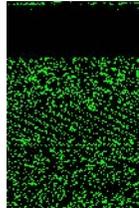
cdex.exe  
(unknown compiler)



apache.exe  
(unknown compiler)



ethereal.exe  
(unknown compiler)



## Comparing Image Files

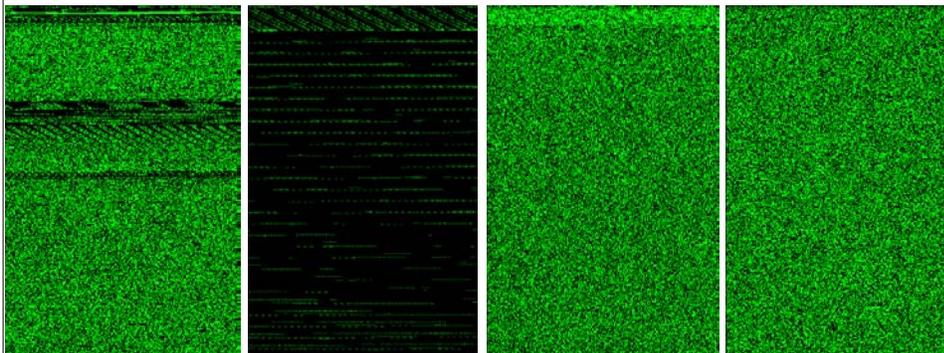
(1 bit per pixel)

image.jpg

image.bmp

image.zip

image.pae  
(encrypted)



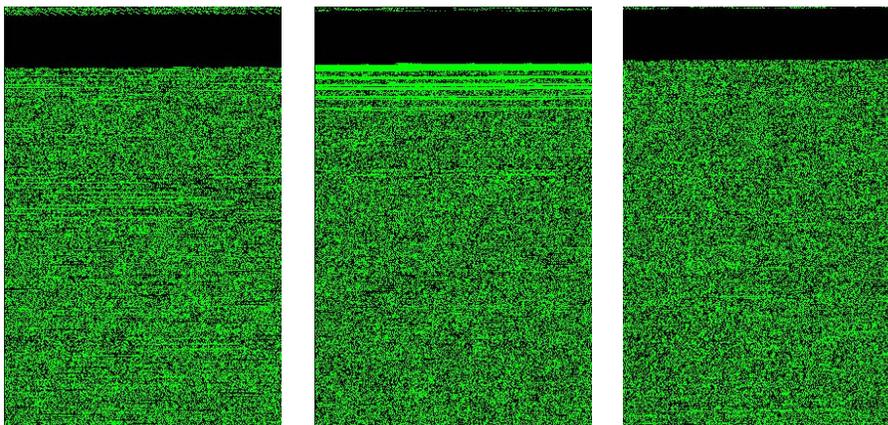
## Comparing mp3 files

(1 bit per pixel)

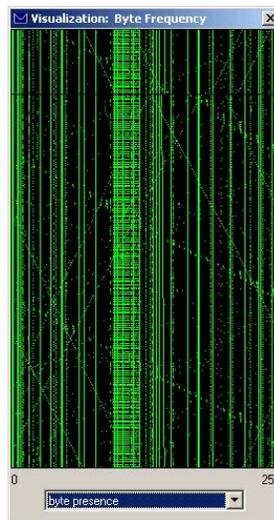
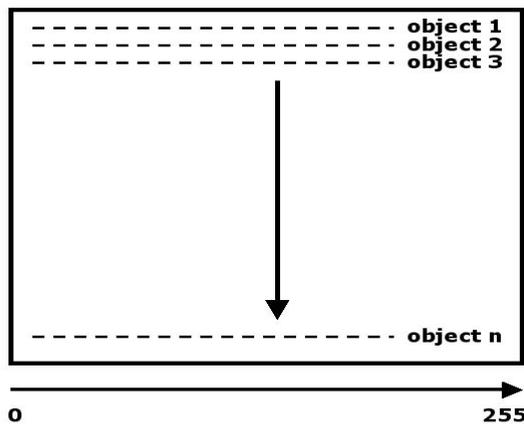
the.mp3

pash.mp3

disguises.mp3



## Byte Visualization

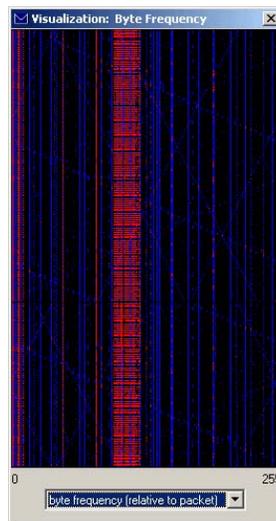


## Byte Presence and Frequency

(lower case dictionary file)

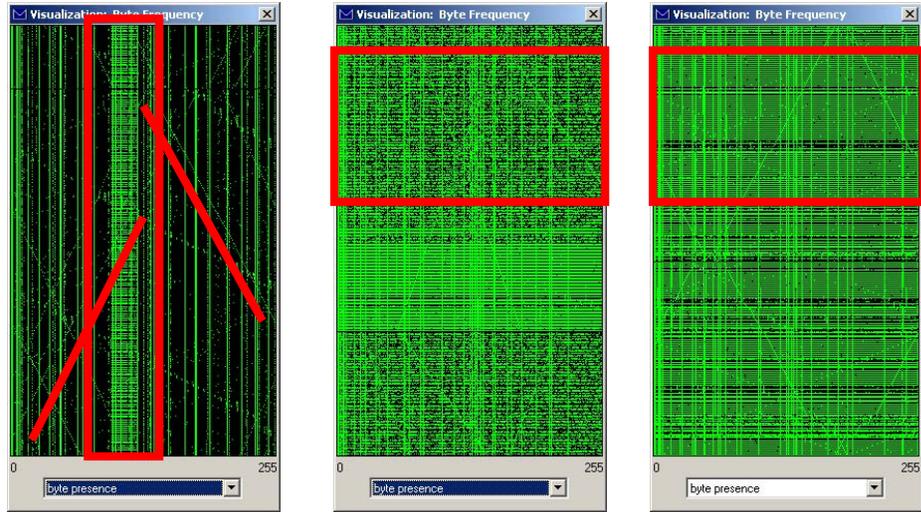


presence



frequency

## Byte Presence

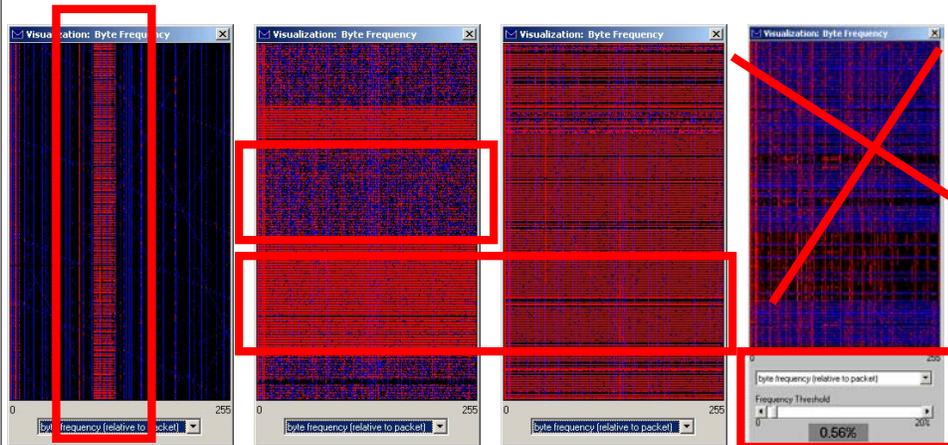


dictionary file via HTTP

ssh

SSL

## Byte Frequency



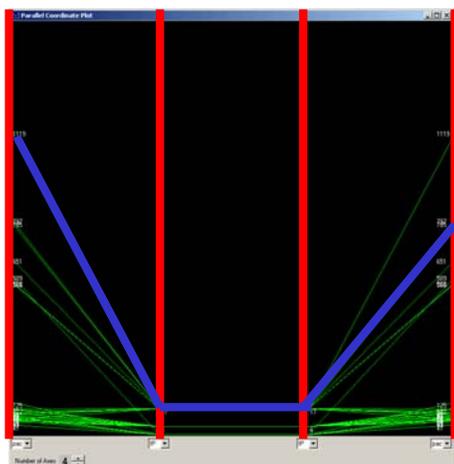
dictionary file over HTTP

ssh

SSL

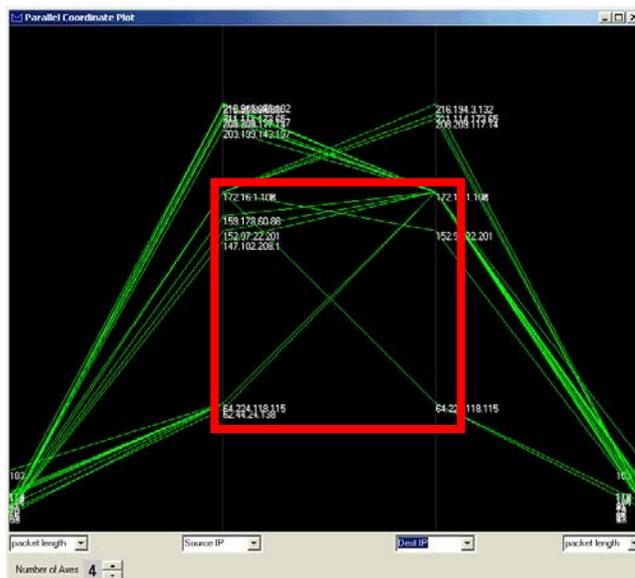
streaming audio

## Parallel Coordinates

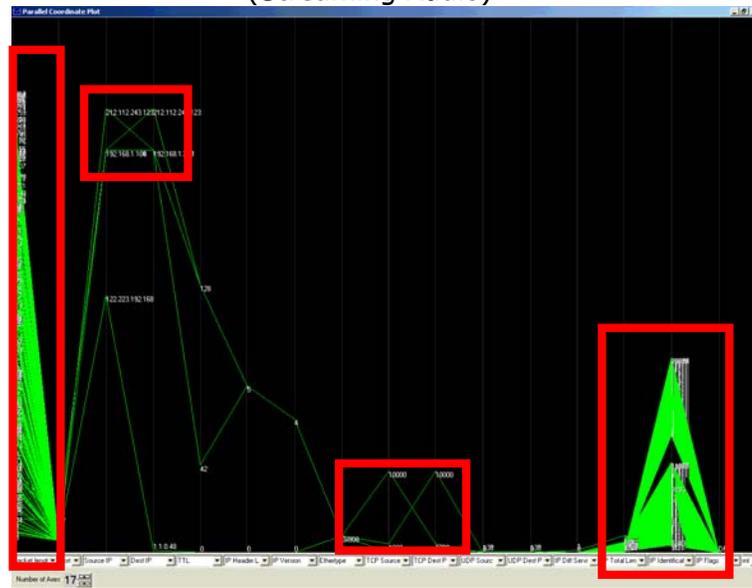


- goal: plot any data fields
- dynamic columns
- change order for different insight
- intelligent lookup and translation of fields
  - e.g. IP transport protocol

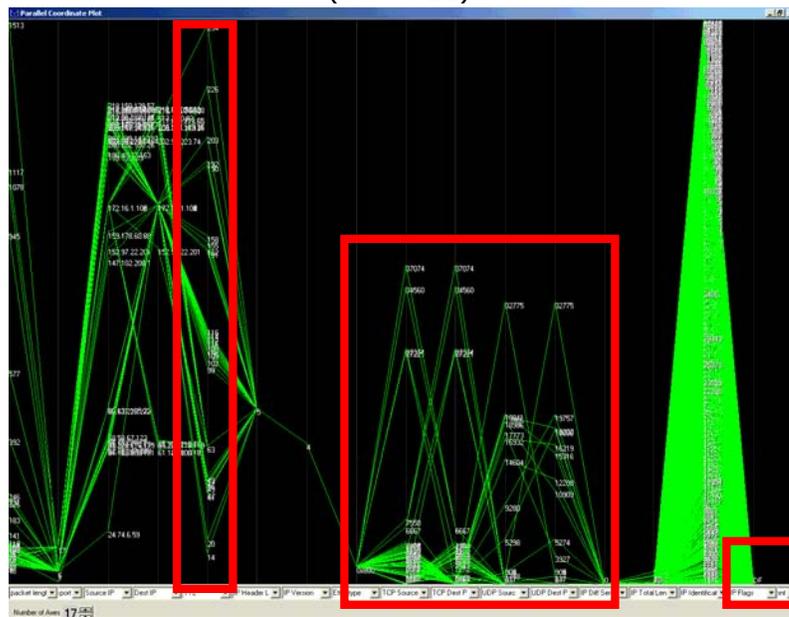
## Parallel Coordinates



### Parallel Coordinates (Streaming Audio)

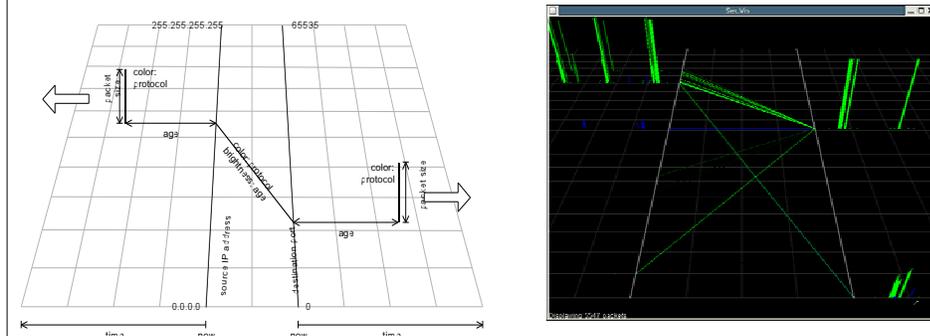


### Parallel Coordinates (SOTM 21)

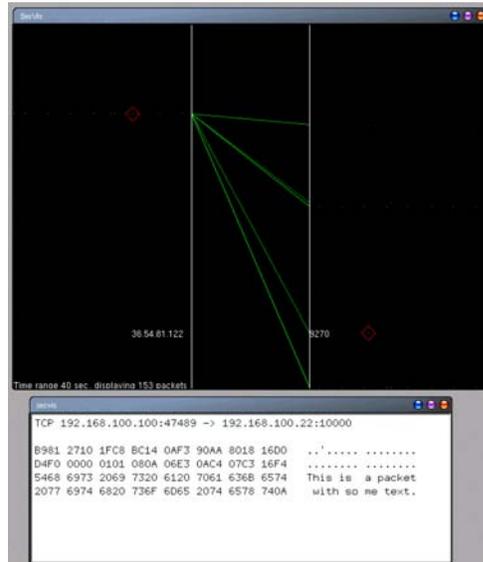




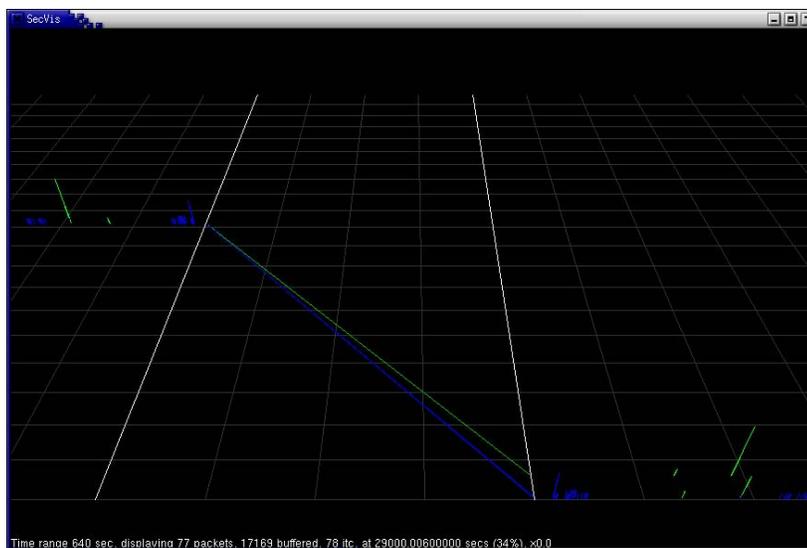
## Krasser Visualization (secvis)



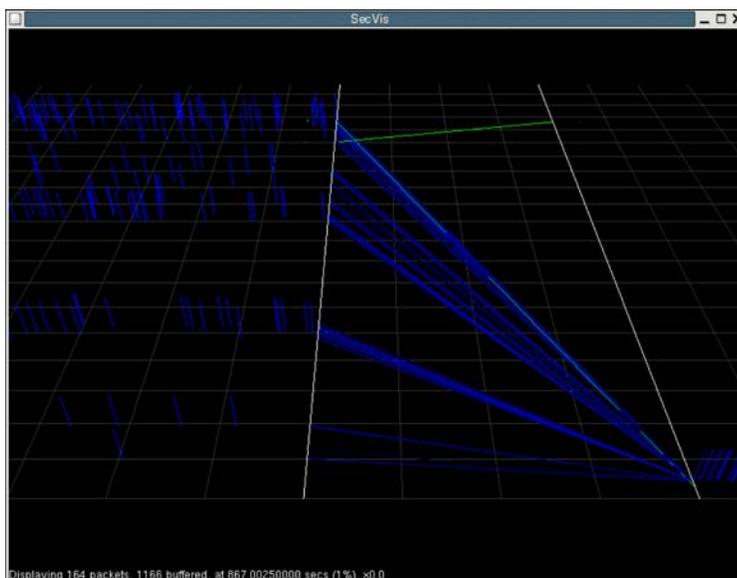
## Overview and Detail



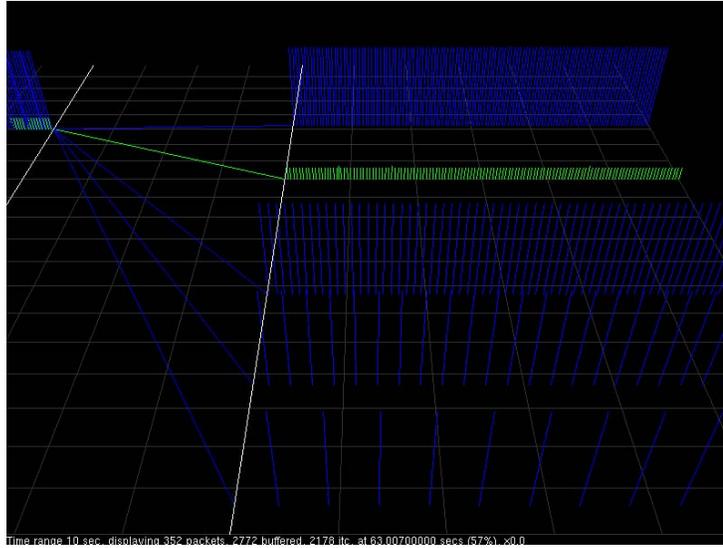
## Routine Honeynet Traffic (baseline)



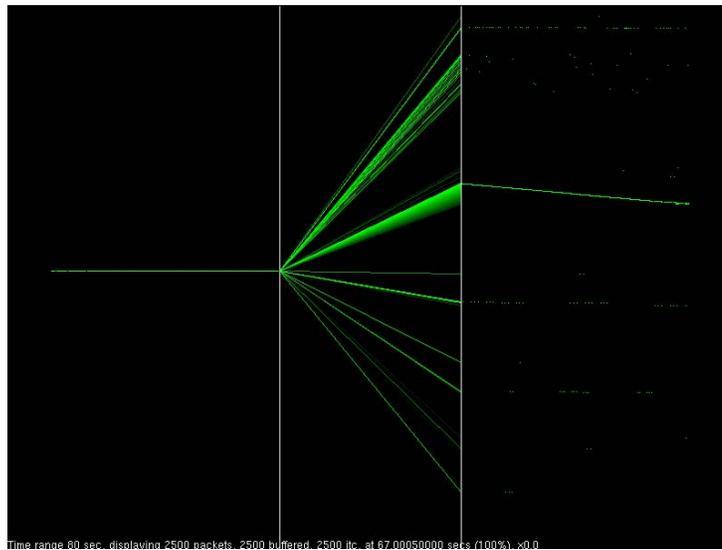
## Slammer Worm



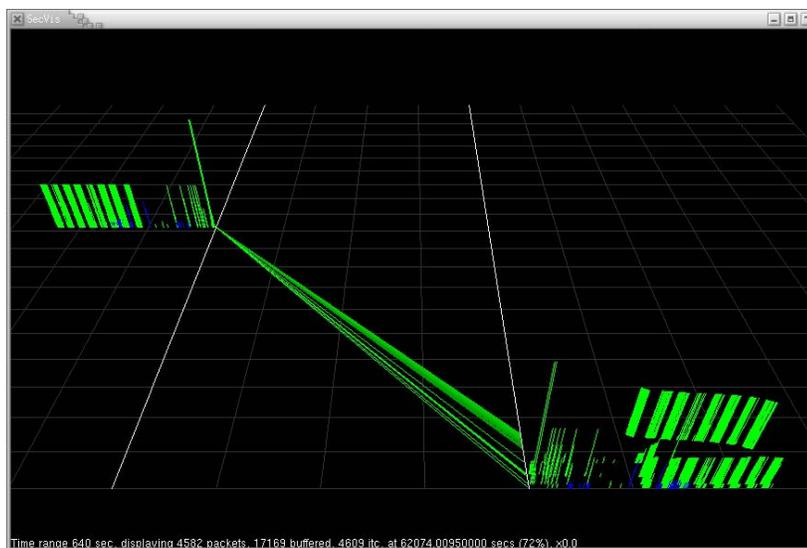
## At a Glance Measurement (Constant Bitrate UDP Traffic)



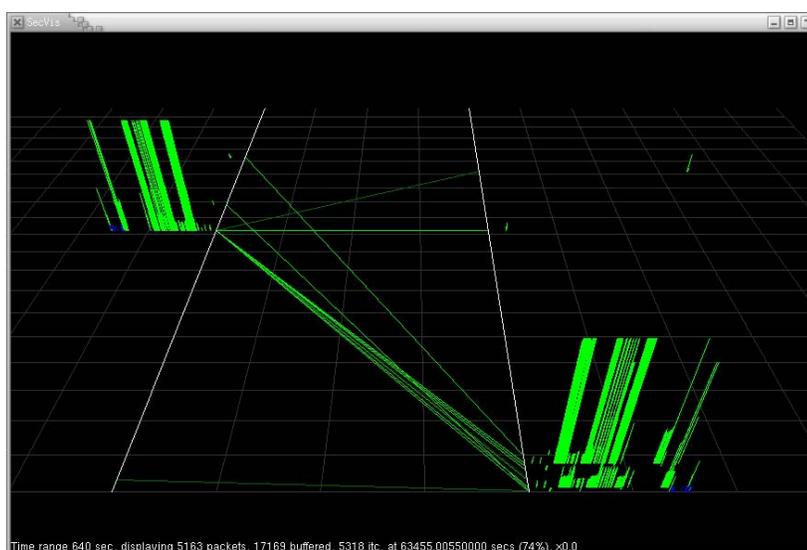
## Port Sweep



## Compromised Honeypot

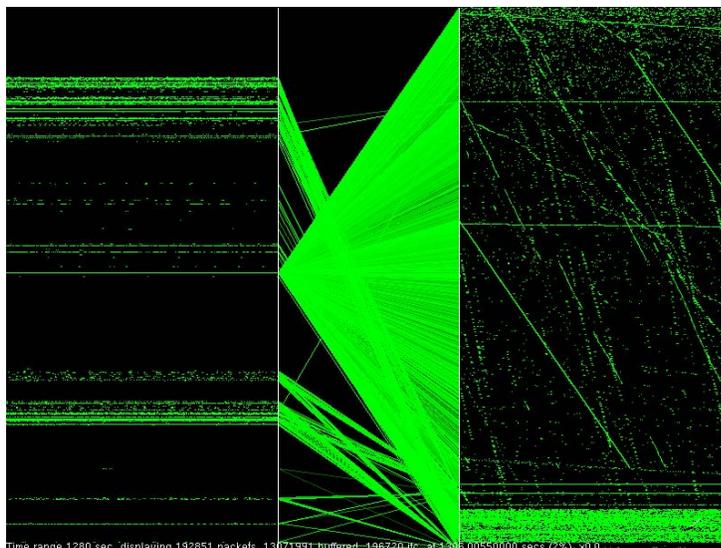


## Attacker Transfers Three Files...





## Combined botnet/honeynet traffic



## For more information...

### Bit Rainfall ([email me...](#))

- G. Conti, J. Grizzard, M. Ahamad and H. Owen; "Visual Exploration of Malicious Network Objects Using Semantic Zoom, Interactive Encoding and Dynamic Queries;" *IEEE Symposium on Information Visualization's Workshop on Visualization for Computer Security (VizSEC)*; October 2005.

### Parallel Coordinate Plots

- Multidimensional Detective by Alfred Inselberg  
<http://www.sims.berkeley.edu/academics/courses/is247/s04/resources/inselberg97.pdf>

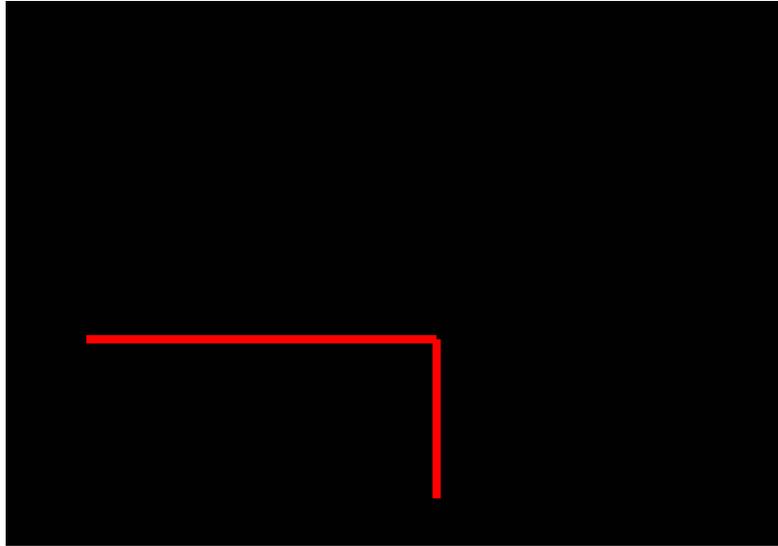
### Byte Frequency Analysis

- Wei-Jen Li, Benjamin Herzog, Ke Wang, Sal Stolfo, "Fileprints: Identifying File Types by N-gram Analysis", *IEEE Information Assurance Workshop*, 2005.
- Ke Wang, Salvatore J. Stolfo. "Anomalous Payload-based Network Intrusion Detection", *Recent Advance in Intrusion Detection (RAID)*, 2004.

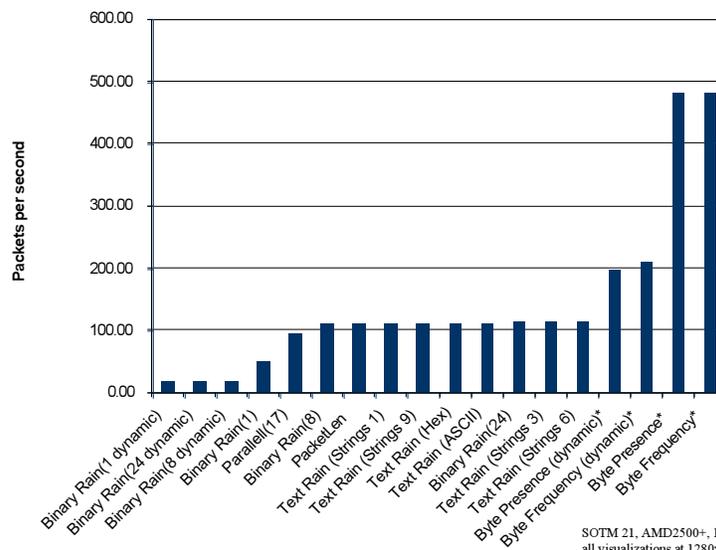
### Krasser Visualization ([see www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti))

- S. Krasser, G. Conti, J. Grizzard, J. Gribshaw and H. Owen; "Real-Time and Forensic Network Data Analysis Using Animated and Coordinated Visualization;" *IEEE Information Assurance Workshop (IAW)*; June 2005.

## Open GL System Performance (secvis)

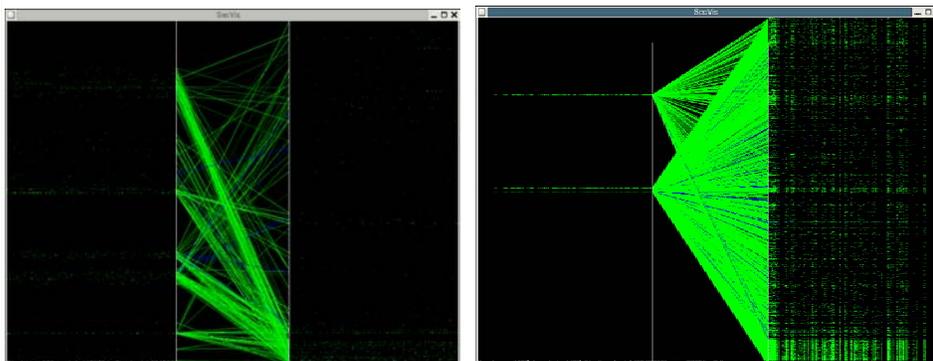


## Win32 Performance (SOTM 21, 3389 packets, rumint v1.60)



SOTM 21, AMD2500+, 1GB RAM  
all visualizations at 1280x1024 except  
byte frequency and presences which are fixed at 256x418

## Campus Network Traffic (10 msec capture)



inbound

outbound

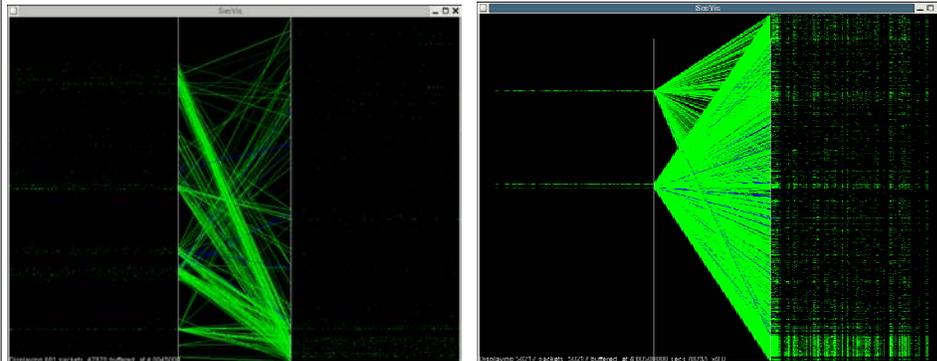
## Directions for the Future...

We are only scratching the surface of the possibilities

- attack specific community needs
- launch network packets?
- protocol specific visualizations
  - including application layer (e.g. VoIP, HTTP)
- Open GL
- graph visualization+
- screensaver/wallpaper snapshot?
- work out GUI issues
- stress testing
- evaluate effectiveness

CTF Visualization (coming soon)

## Campus Network Traffic (10 msec capture)



inbound

outbound

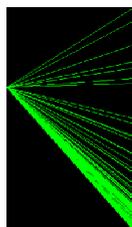
## Directions for the Future...

We are only scratching the surface of the possibilities

- attack specific community needs
- launch network packets?
- protocol specific visualizations
  - including application layer (e.g. VoIP, HTTP)
- Open GL
- graph visualization+
- screensaver/wallpaper snapshot?
- work out GUI issues
- stress testing
- evaluate effectiveness

CTF Visualization (coming soon)

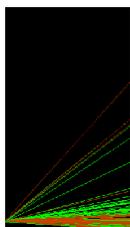
## Library of Tool Fingerprints



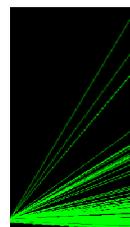
nmap 3 (RH8)



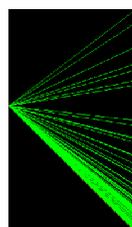
nmap 3 UDP (RH8)



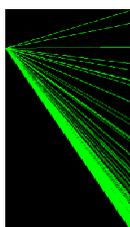
scanline 1.01 (XP)



SuperScan 3.0 (XP)



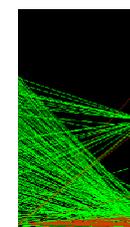
NMapWin 3 (XP)



nmap 3.5 (XP)



nikto 1.32 (XP)



SuperScan 4.0 (XP)

## For more information...

G. Conti and K. Abdullah; " Passive Visual Fingerprinting of Network Attack Tools;" ACM Conference on Computer and Communications Security's Workshop on Visualization and Data Mining for Computer Security (VizSEC); October 2004.

--Talk PPT Slides

G. Conti; "Network Attack Visualization;" DEFCON 12; August 2004.

--Talk PPT Slides  
 --Classical InfoVis Survey PPT Slides  
 --Security InfoVis Survey PPT Slides

see [www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti)

# Future Interaction



- Database of filters, snapshots and analyst mark-up (shareable)
- Both command line and GUI filtering
- Network response

[http://www.smsu.edu/etc/images/photo\\_mixingboard.jpg](http://www.smsu.edu/etc/images/photo_mixingboard.jpg)

<http://depts.washington.edu/sag/services/workshops/multimedia/daw/img/8.jpg>

# Attacking the Analyst...

G. Conti, M. Ahamad and J. Stasko;  
 "Attacking Information Visualization System Usability: Overloading and Deceiving the Human;" *Symposium on Usable Privacy and Security (SOUPS)*; July 2005. **On the CD...**

G. Conti and M. Ahamad; "A Taxonomy and Framework for Countering Denial of Information Attacks;" *IEEE Security and Privacy*. (accepted, to be published) **Email me...**

DEFCON CTF **DoI vs. DOS...**



## On the CD...

- Code
  - rumint
  - secvis
  - rumint file conversion tool (pcap to rumint)
- Papers
  - SOUPS Malicious Visualization paper
  - Hacker conventions article
- Data
  - SOTM 21 .rum



See also: [www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti) and [www.rumint.org](http://www.rumint.org)

## Feedback Requested...

- Tasks
- Usage
  - provide feedback on GUI
  - needed improvements
  - multiple monitor machines
  - performance under stress
  - bug reports
- Data
  - interesting packet traces
  - screenshots
    - with supporting .rum and .pcap files, if possible
- Pointers to interesting related tools (viz or not)
- New viz and other analysis ideas

**Volunteers to participate in user study**

## Acknowledgements

404.se2600, Kulsoom Abdullah, Sandip Agarwala, Mustaque Ahamad, Bill Cheswick, Chad, Clint, Tom Cross, David Dagon, DEFCON, Ron Dodge, EliO, Emma, Mr. Fuzzy, Jeff Gribshaw, Julian Grizzard, GTISC, Hacker Japan, Mike Hamelin, Hendrick, HoneyNet Project, Interz0ne, Jinsuk Jun, Kenshoto, Oleg Kolesnikov, Sven Krasser, Chris Lee, Wenke Lee, John Levine, David Maynor, Jeff Moss, NETI@home, Henry Owen, Dan Ragsdale, Rockit, Byung-Uk Roho, Charles Robert Simpson, Ashish Soni, SOUPS, Jason Spence, John Stasko, Strick, Susan, USMA ITOC, IEEE IAW, VizSEC 2004, Grant Wagner and the Yak.



- 100+ Graduate Level InfoSec Researchers
- Multiple InfoSec degree and certificate programs
- Representative Research
  - User-centric Security
  - Adaptive Intrusion Detection Models
  - Defensive Measures Against Network Denial of Service Attacks
  - Exploring the Power of Safe Areas of Computation
  - Denial of Information Attacks (Semantic Hacking)
  - Enterprise Information Security
- **Looking for new strategic partners, particularly in industry and government**

[www.gtisc.gatech.edu](http://www.gtisc.gatech.edu)

Questions?

**Greg Conti**  
conti@cc.gatech.edu

[www.cc.gatech.edu/~conti](http://www.cc.gatech.edu/~conti)  
[www.rumint.org](http://www.rumint.org)

