

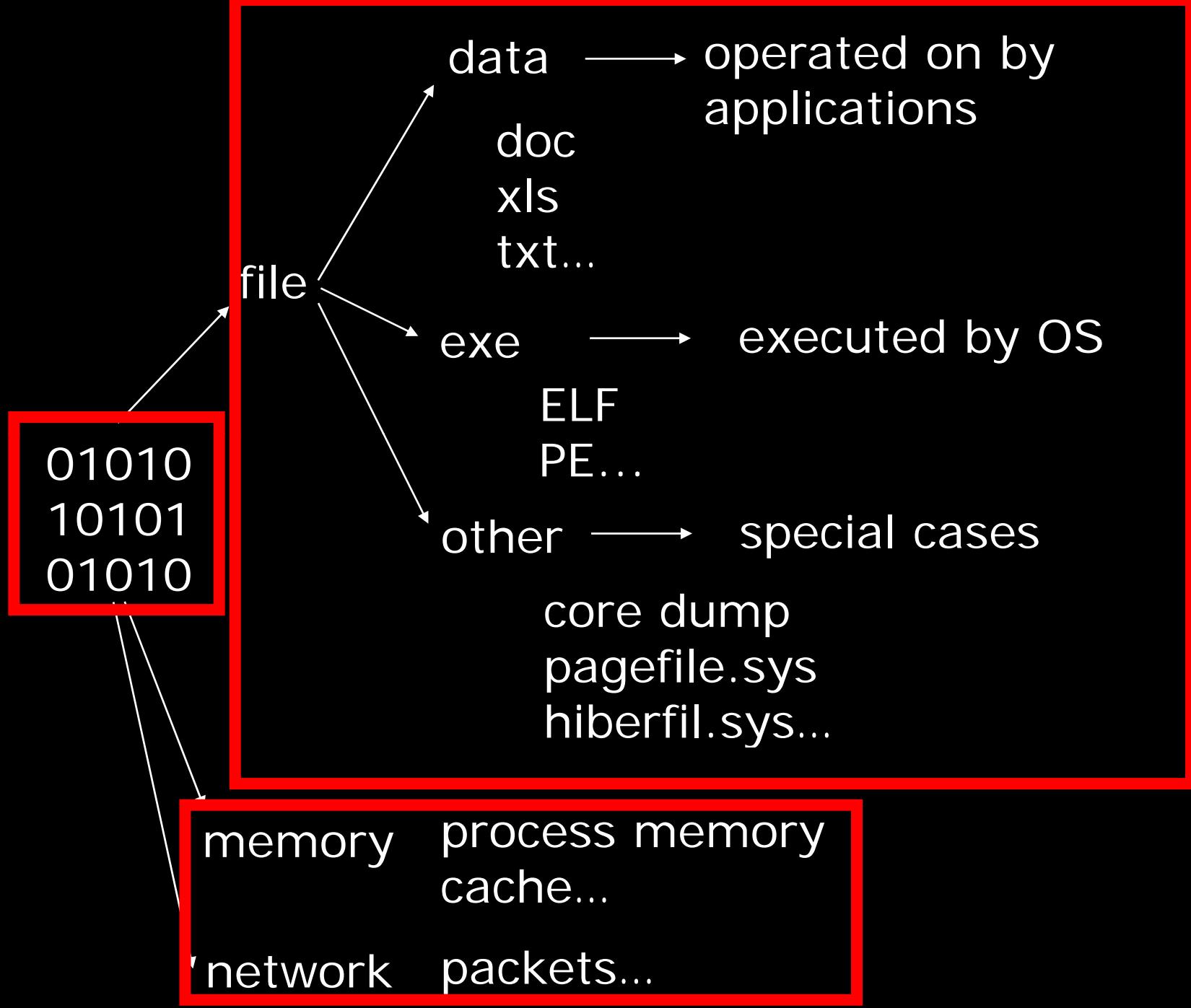
Visual Forensic Analysis and Reverse Engineering of Binary Data

*Gregory Conti
Erik Dean*

*United States Military Academy
West Point, New York
gregory.conti@usma.edu
erik.dean@usma.edu*

Outline

- The Problem – Tiny Windows
- Background and Motivation
- Related Work
- Moving Beyond Hex
- System Design
- Case Studies
- Demos



high
insight

Ida Pro
OllyDBG
BinNavi (Zynamics)
BinDiff (Zynamics)...

Filemon
Regmon...

lower
insight

011

hex editors
hexdump
grep & diff
strings

objdump
original
application

general purpose

precise application

strings /grep/diff

```
H:\Datasets>strings 20040517_homeISP.pcap | more
```

Strings v2.4

Copyright (C) 1999-2007 Mark Russinovich

Sysinternals - www.sysinternals.com

0hF

M@Y

7bs

Z19Z

MICROSOFT NETWORKS

WINDOWS USER

Microsoft Security Bulletin MS03-043

Buffer Overrun in Messenger Service Could Allow Code Execution
(828035)

Affected Software:

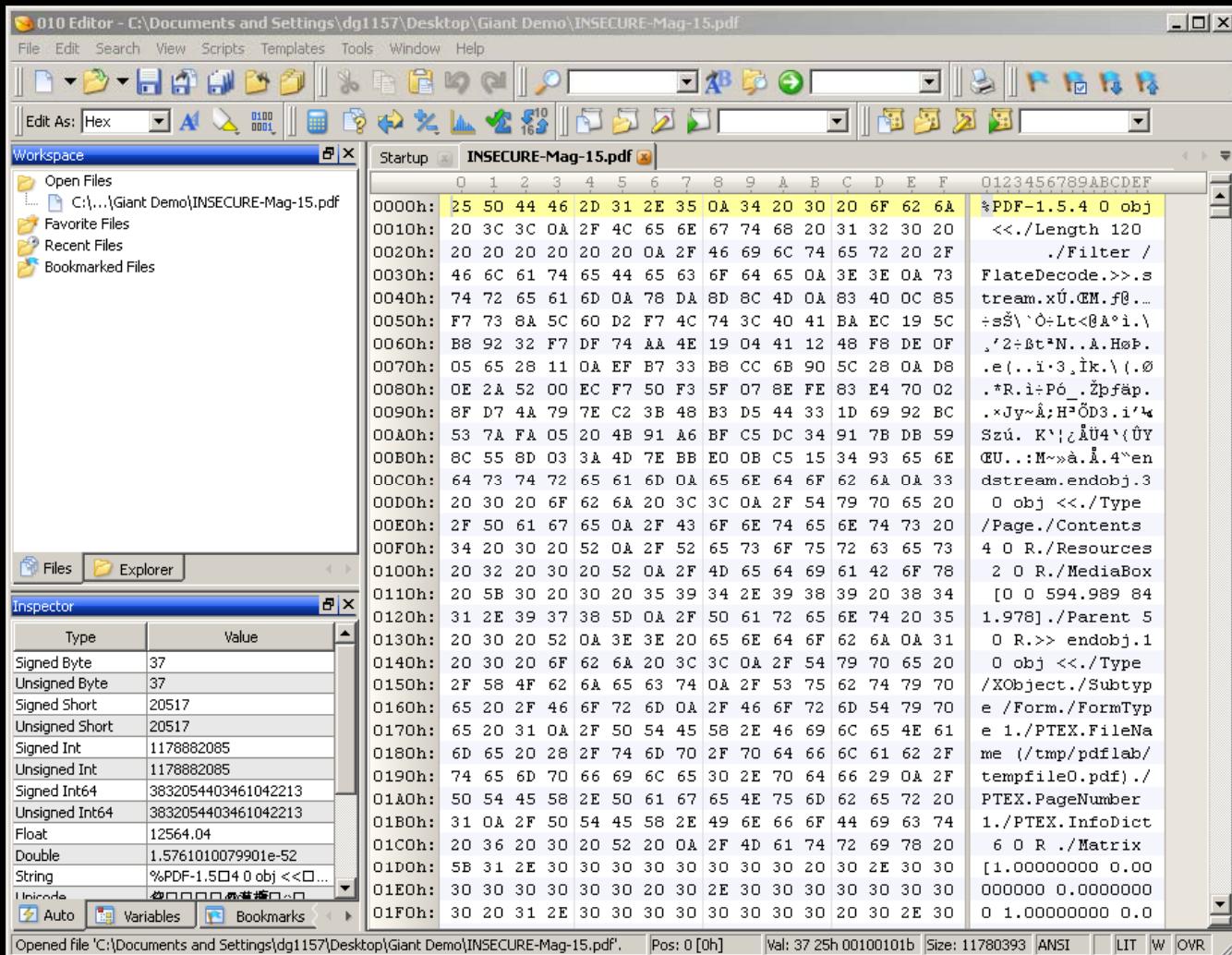
Microsoft Windows NT Workstation

Microsoft Windows NT Server 4.0

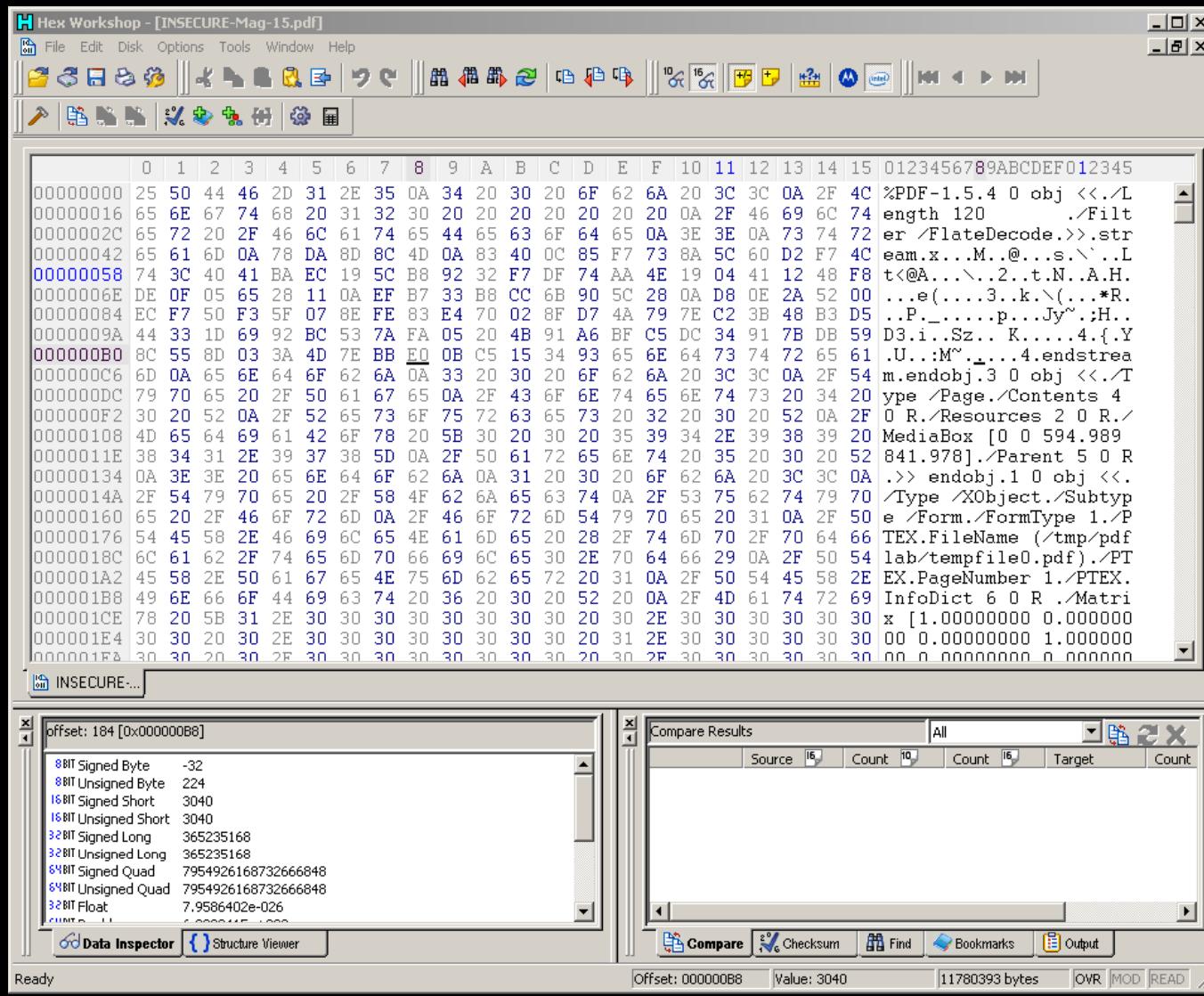
Microsoft Windows 2000

...

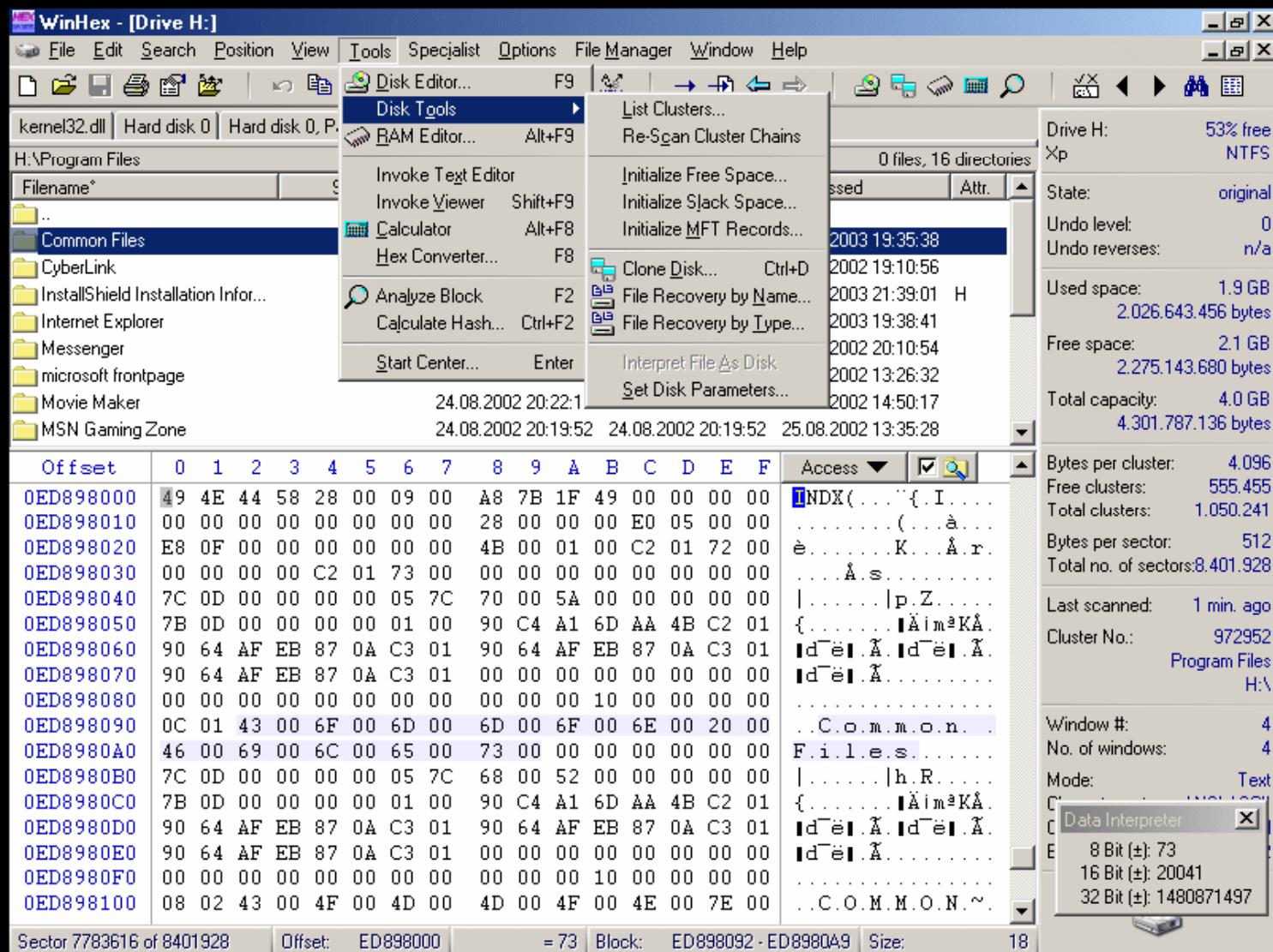
011 Hex Editor



Hex Workshop



WinHex



high
insight

Ida Pro
OllyDBG
BinNavi (Zynamics)
BinDiff (Zynamics)...

Filemon
Regmon...

lower
insight

011

hex editors
hexdump
grep & diff
strings

objdump
original
application

general purpose

precise application

high
insight

011

lower
insight

hex editors
hexdump
grep & diff
strings

general purpose

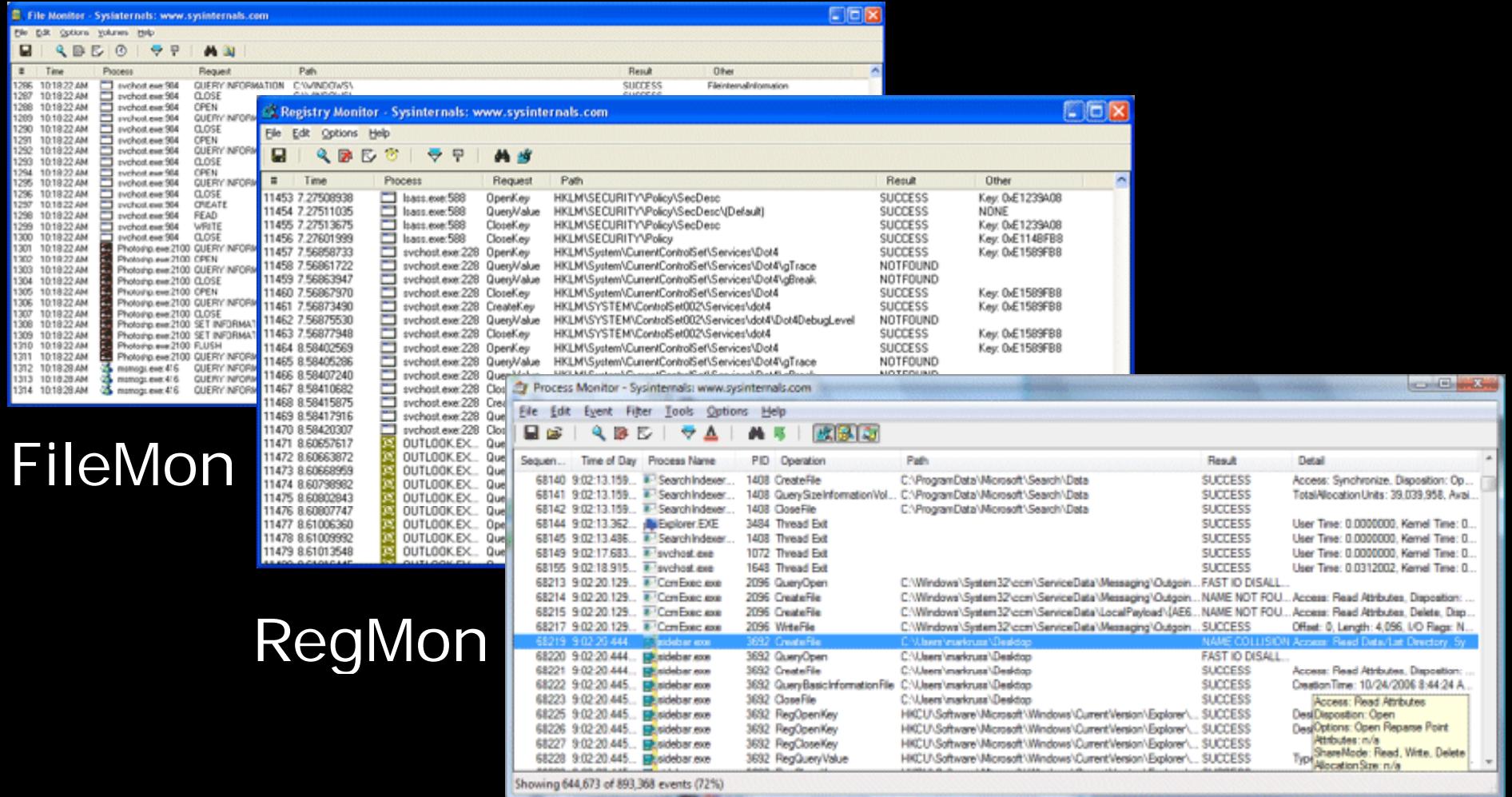
Ida Pro
OllyDBG
BinNavi (Zynamics)
BinDiff (Zynamics)...

Filemon
Regmon...

objdump
original
application

precise application

SysInternals



Process Monitor

• • •

<http://technet.microsoft.com/en-us/sysinternals/default.aspx>

Wireshark

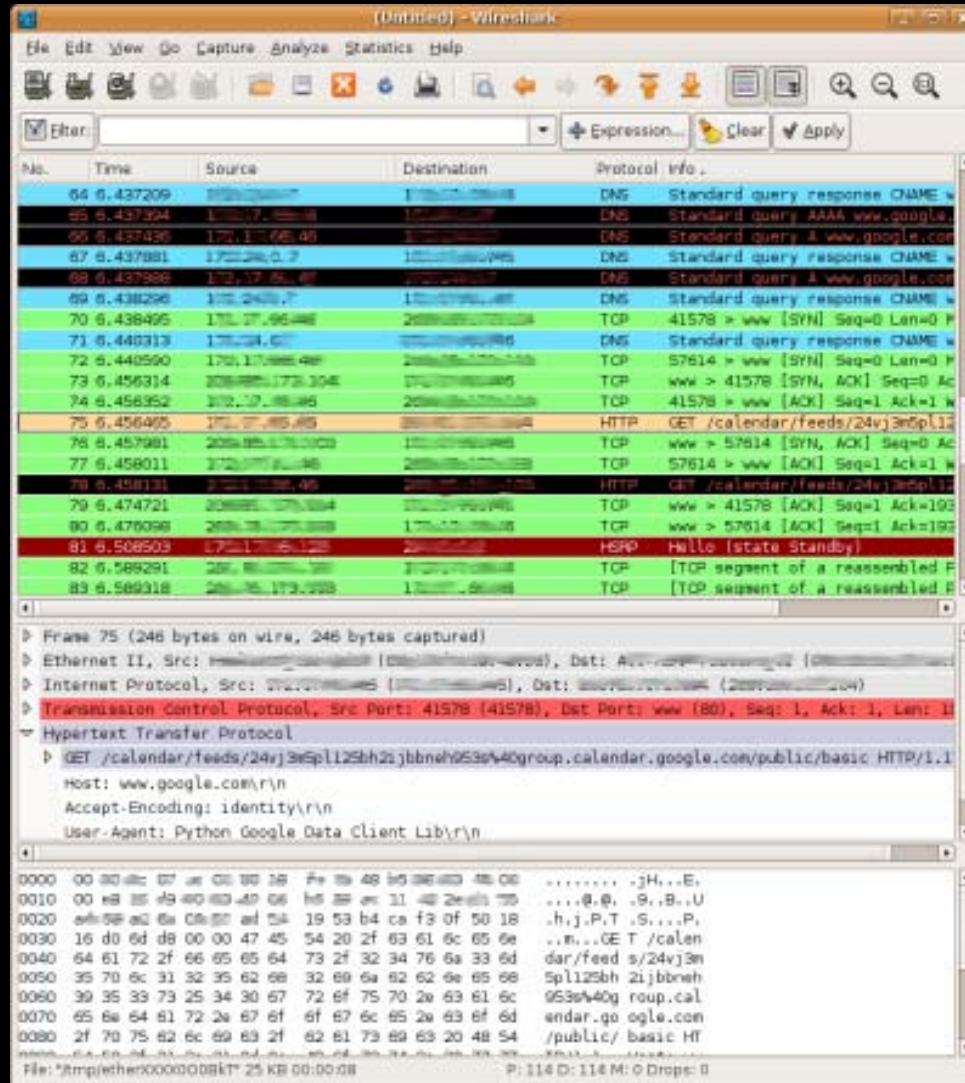
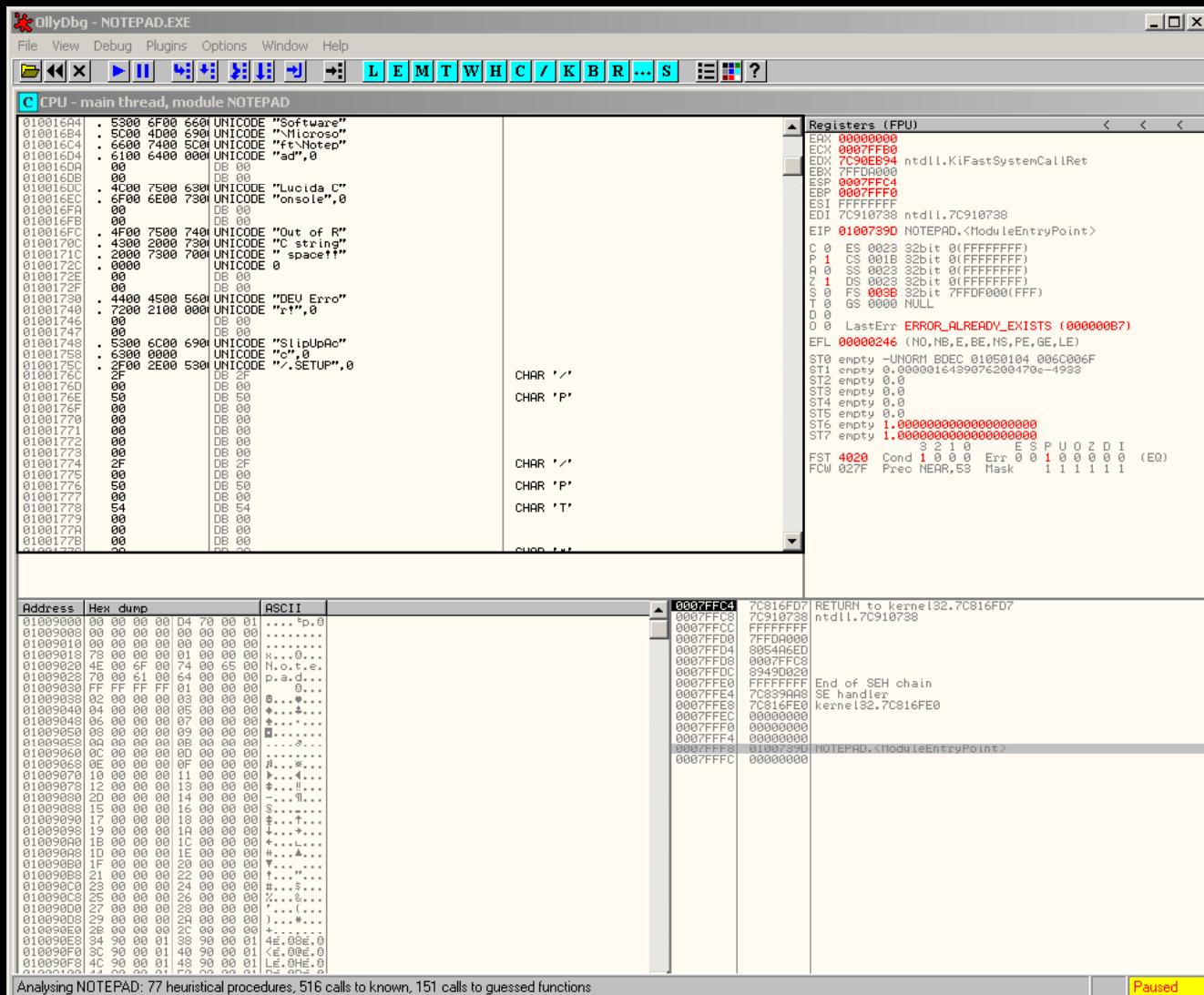


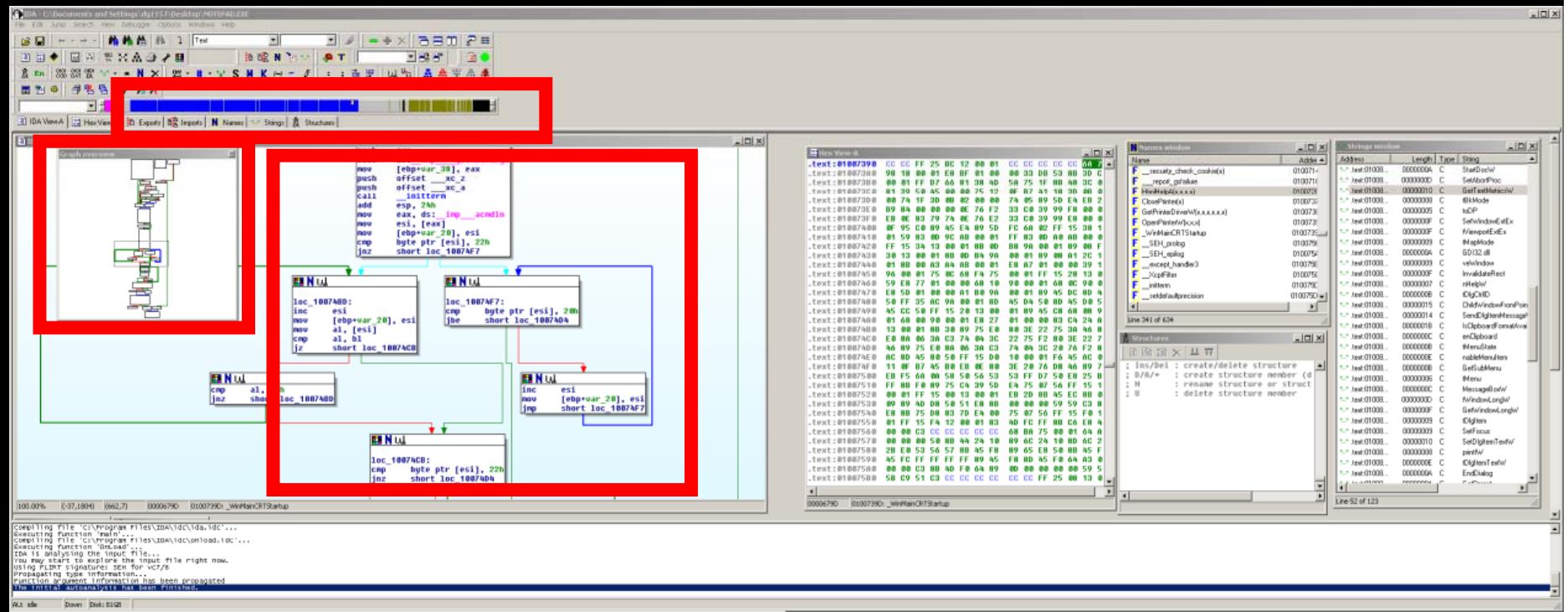
image: <http://code.google.com/support/bin/answer.py?answer=71567>

OllyDbg



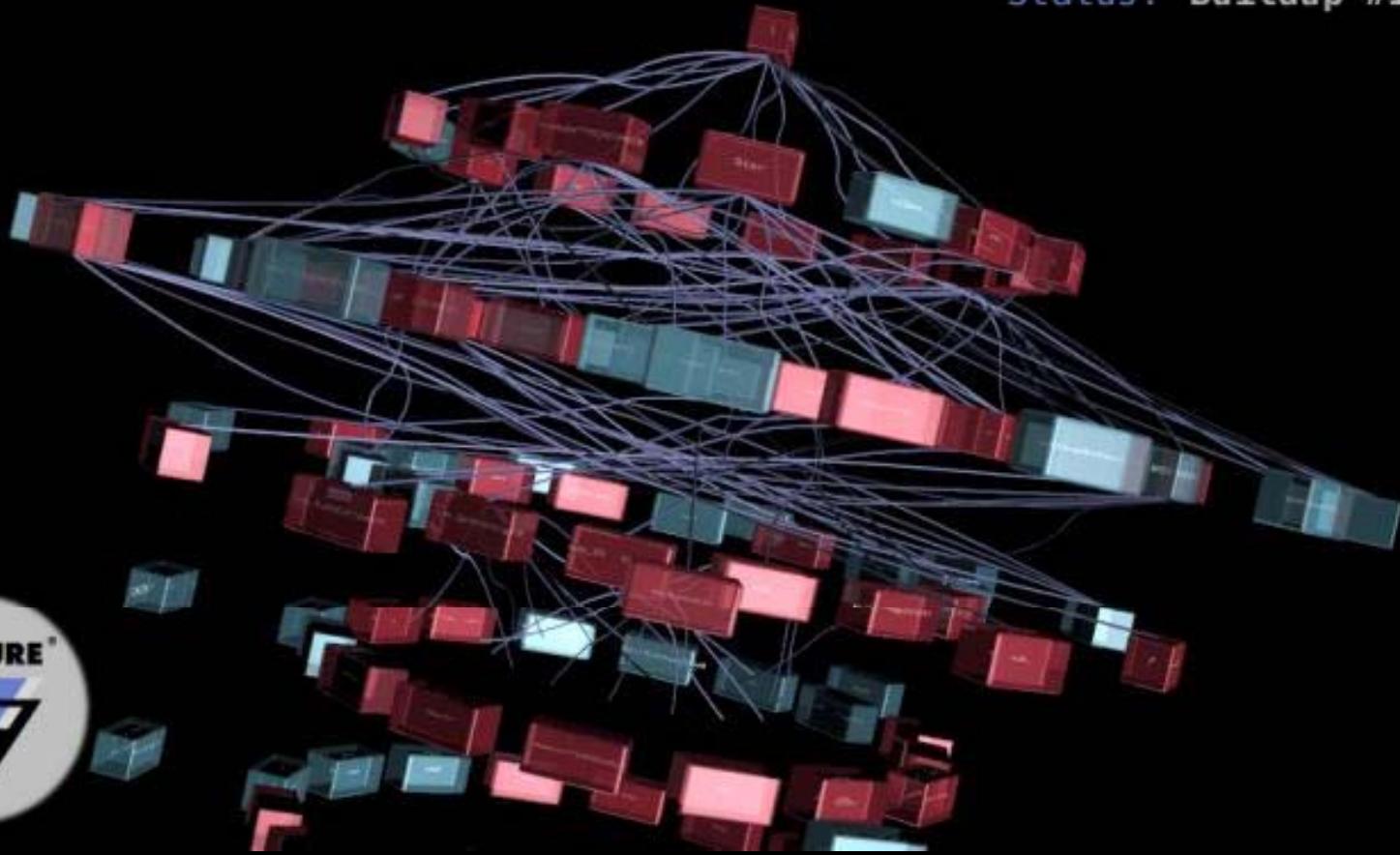
IDA Pro

v5.1



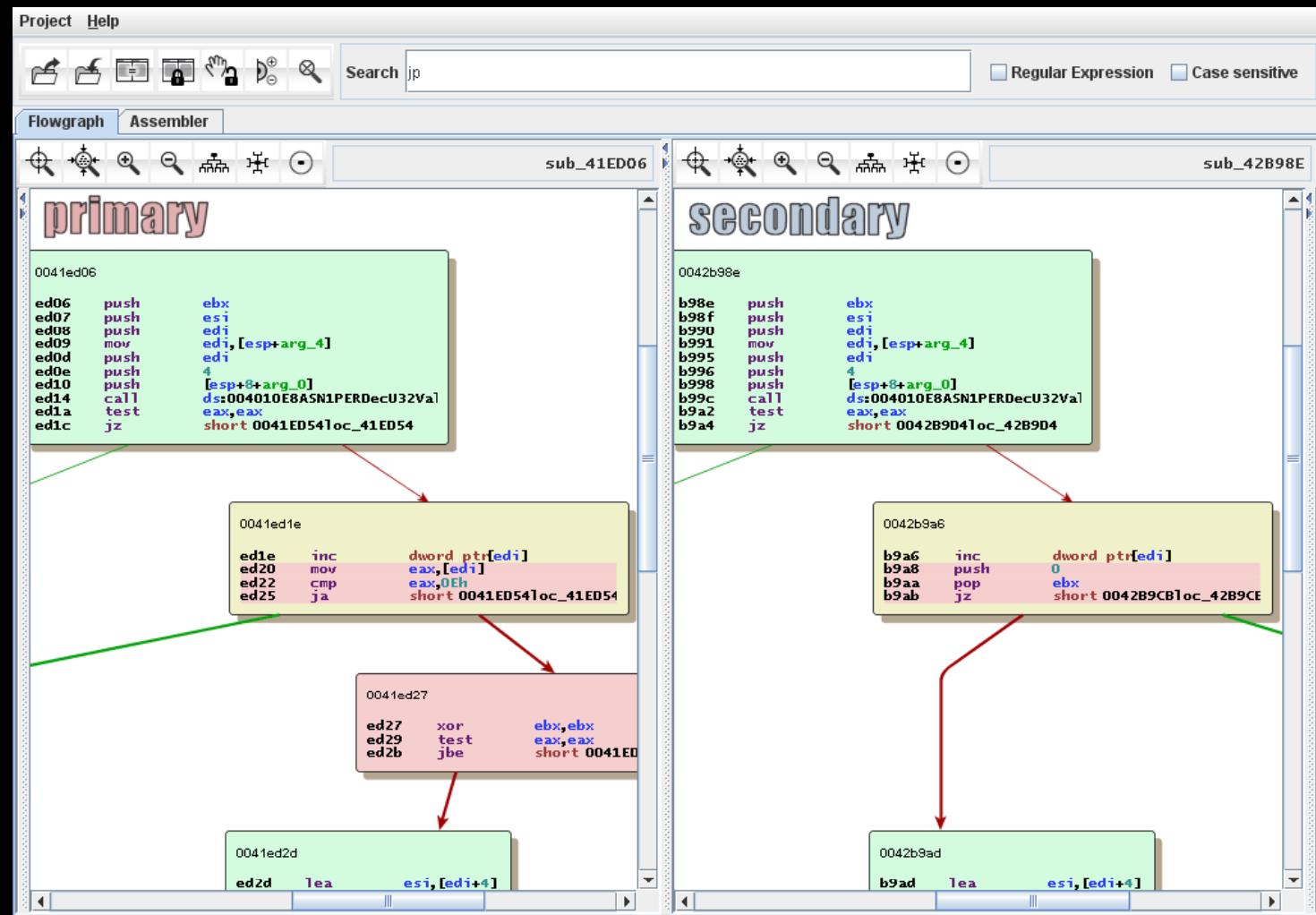
F-Secure Malware

Sample: W32/Bagle.AG@mm
Status: Buildup #1

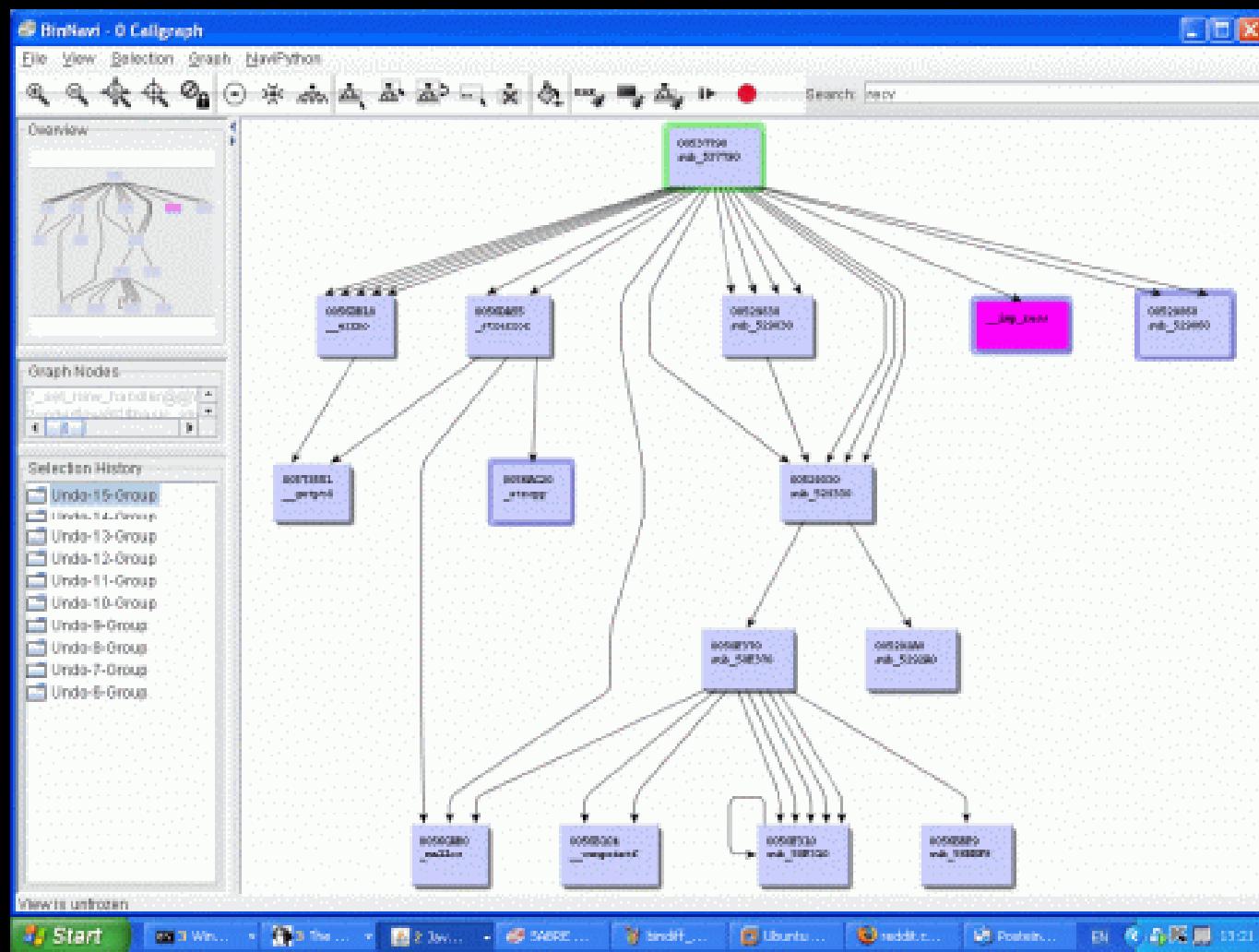


<http://www.f-secure.com/weblog/archives/00000662.html>

Zynamics BinDiff



Zynamics BinNavi



<http://www.zynamics.com/index.php?page=binnavi>

high
insight

011

lower
insight

hex editors
hexdump
grep & diff
strings

general purpose

Ida Pro
OllyDBG
BinNavi (Zynamics)
BinDiff (Zynamics)...

Filemon
Regmon...

objdump
original
application

precise application

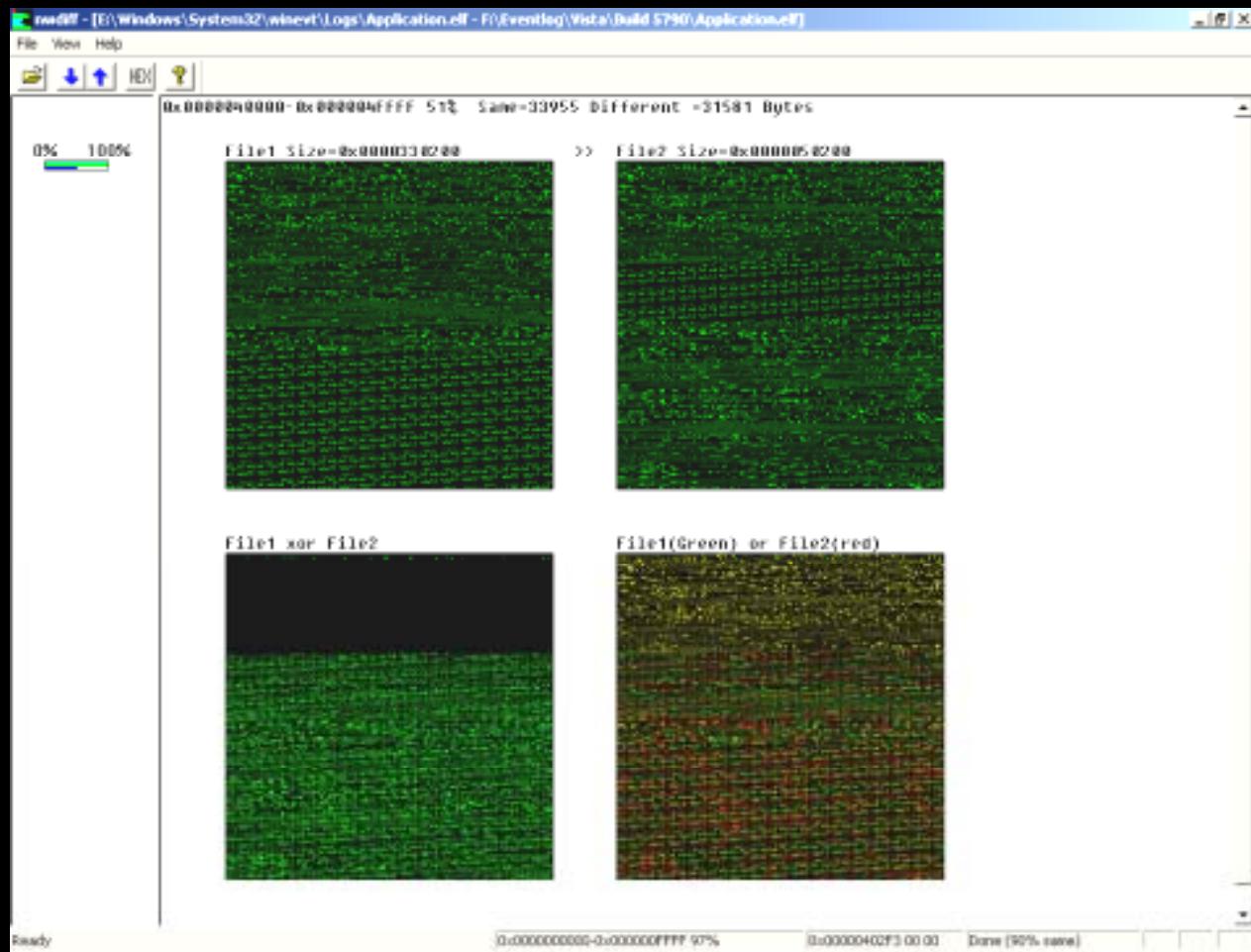
Framework

- File Independent Level
 - Entropy
 - Byte Frequency
 - N-Gram Analysis
 - Strings
 - Hex / Decimal / ASCII
 - Bit Plot (2D/3D)
 - File Statistics
- File Specific Level
 - Complete or Partial Knowledge of File Structure
 - For Example, Metadata

Syntax Highlighting for Hex Dumps (Kaminsky)

image: Dan Kaminsky, CCC2006

nwdiff

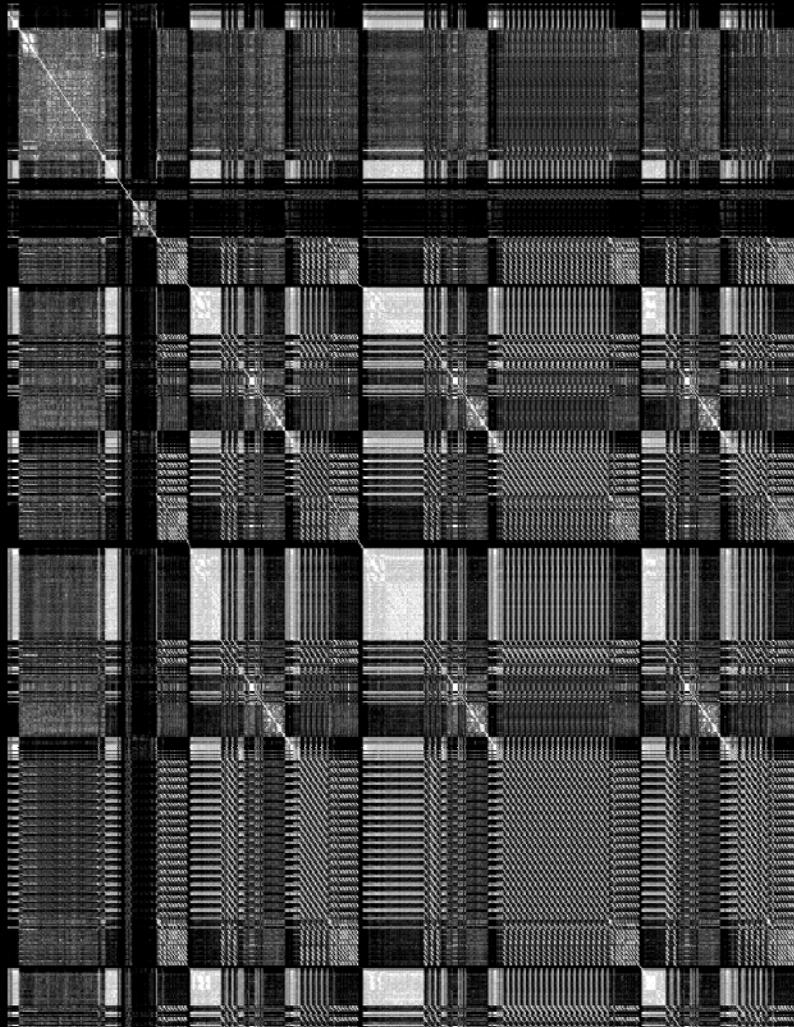


http://computer.forensikblog.de/en/2006/02/compare_binary_files_with_nwdiff.html

http://www.geocities.jp/belden_dr/ToolNwdiff_Eng.html

Dot Plots & Visual BinDiff

(Kaminsky)



Self-Similarity in
a single file. (.NET Assembly)



Diffing Two Files

images: Dan Kaminsky, CCC2006

Textual
Hex/ASCII
Detail View

Traditional
Textual
Utilities
(strings...)

Graphical
Displays

Machine Assisted Mapping and Navigation

Hex Editor Core

Towards a Visual Hex Editor

- Identify Unknown Binaries
- Malware Analysis
- Analyze Unknown/Undocumented File Format
- Locate Embedded Objects
 - Encoding / Encryption
- Audit Files for Vulnerabilities
- Compare files (Diffing)
- Cracking
- Cryptanalysis
- Perform Forensic Analysis
- File System Analysis
- Reporting
- File Fuzzing

Goals

- Handle Large Files
- Many Insightful Windows
- Big Picture Context
- Improved Navigation
- Data Files / Executable Files
- Hex Editor best practices is the foundation
- Support Art & Science

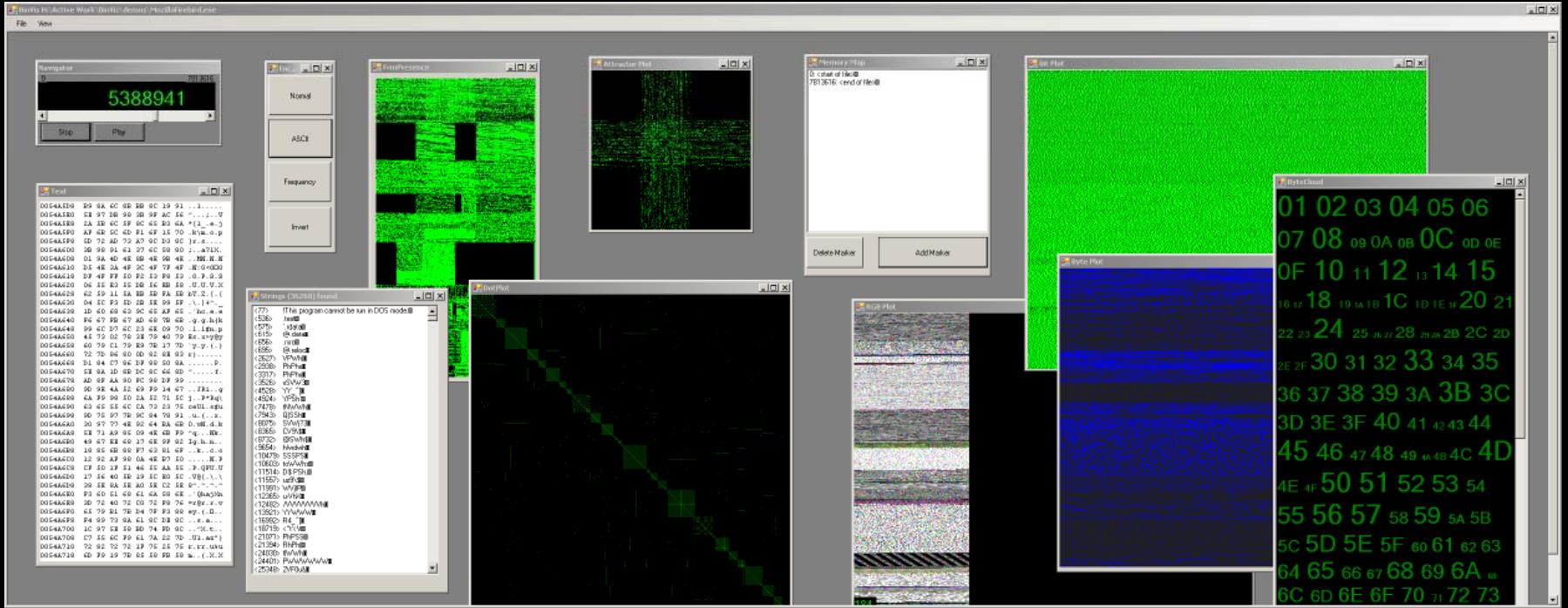
Design

- Robust extensible framework
- Open source
- Context Independent File Analysis
- Semantic File Analysis
- Useful
- Multiple coordinated views
- Combine Functionality of current tools and extend with visuals

Filtering + Encoding

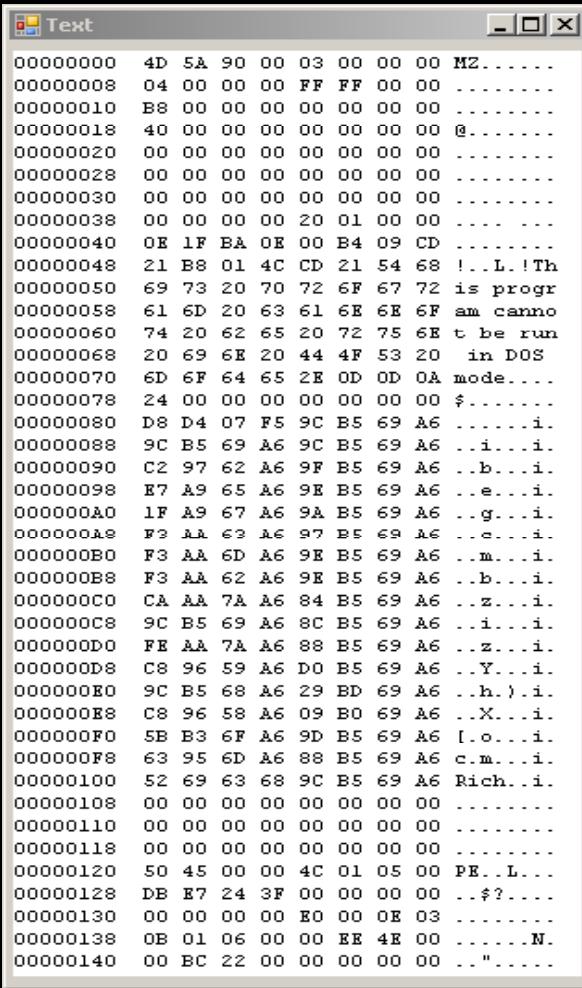
- Identifying something
 - REGEX → algorithmic
- Using this knowledge to..
 - highlight
 - fade
 - remove
- Interactive or automated





- **Textual**: Text/ASCII, Strings, ByteCloud
- **Graphical**: Bitplot, BytePlot, RGBPlot, BytePresence, ByteFrequency, Digram, Dotplot
- **Interaction**: VCR, Memory Map, Color Coding

Traditional Views



Hex / ASCII View



Strings

Strange Attractors and TCP/IP Sequence Number Analysis

(Michał Zalewski)



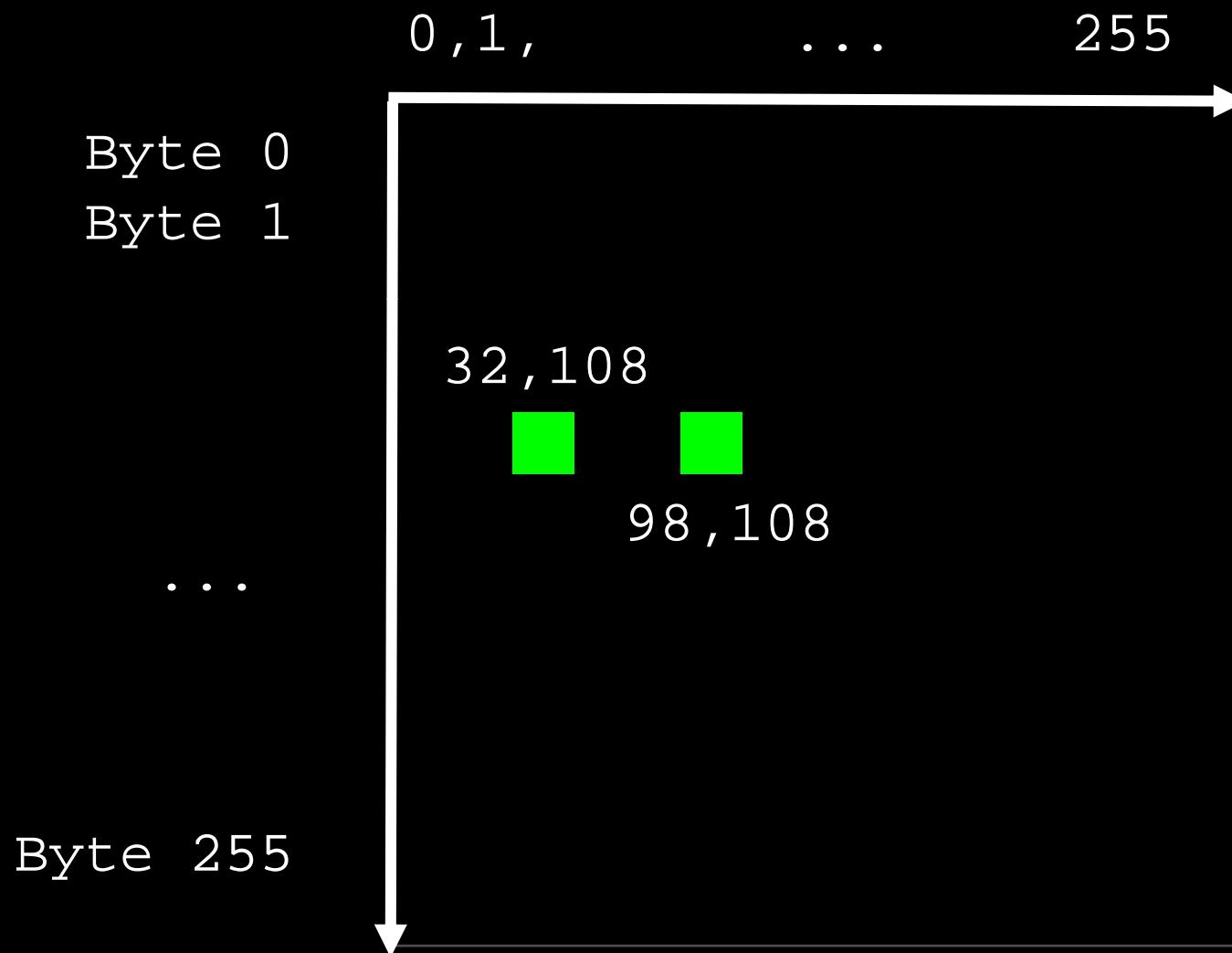
- <http://lcamtuf.coredump.cx/oldtcp/tcpseq.html>
- <http://lcamtuf.coredump.cx/newtcp/>

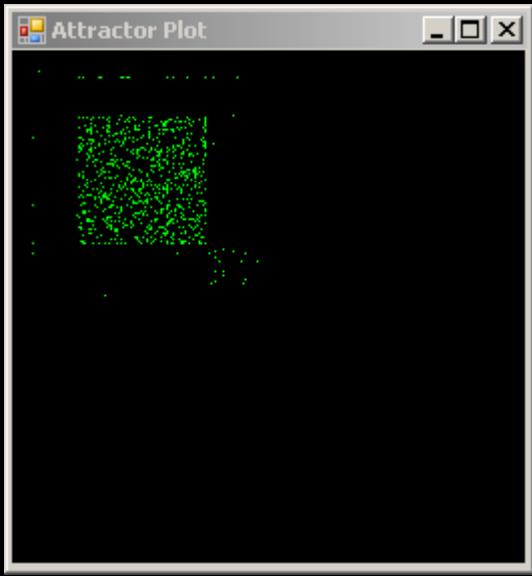
Digraph View

black hat

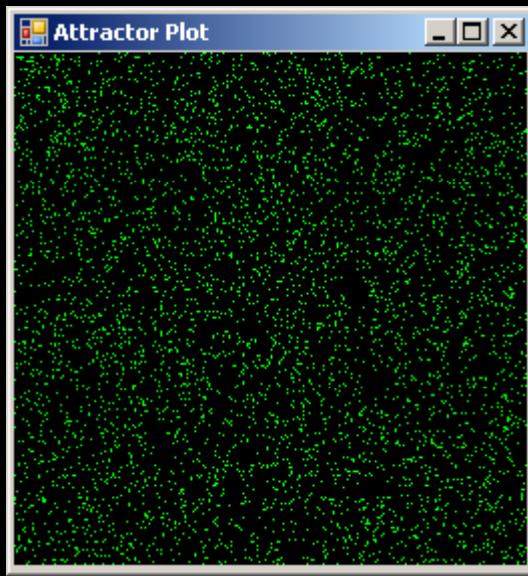
bl	(98 , 108)
la	(108 , 97)
ac	(97 , 99)
ck	(99 , 107)
k_	(107 , 32)
_h	(32 , 104)
ha	(104 , 97)
at	(97 , 116)

Digraph View

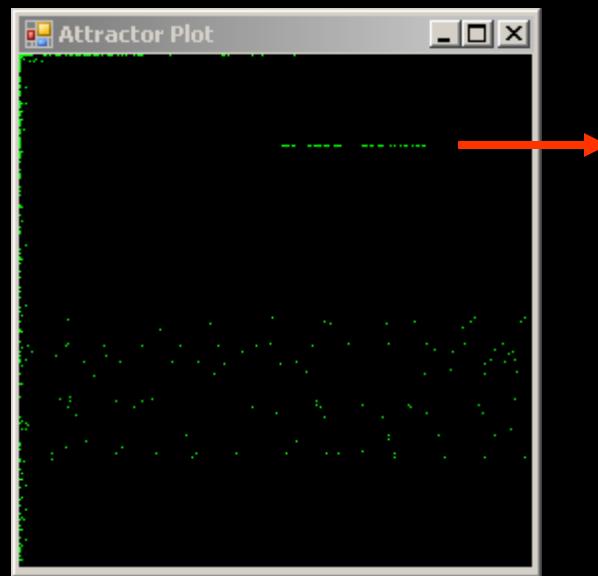




uuencoded



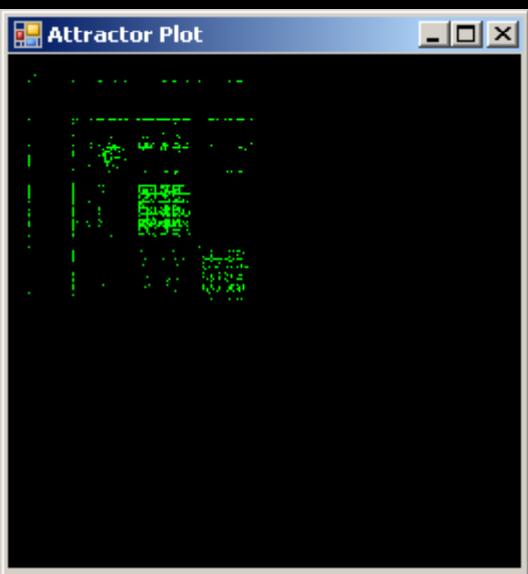
compression
encryption



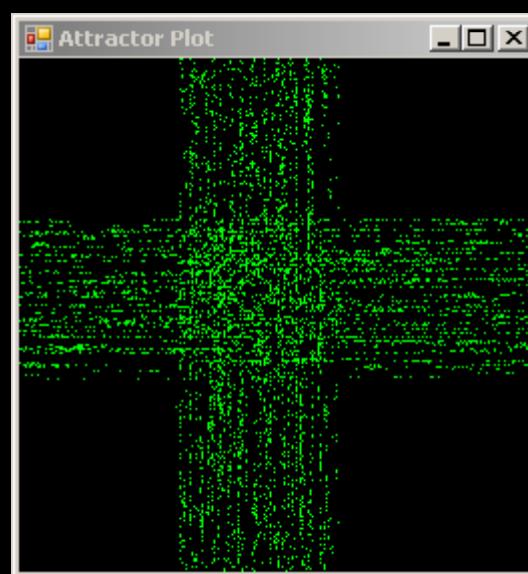
incrementing
words



slashdot.org

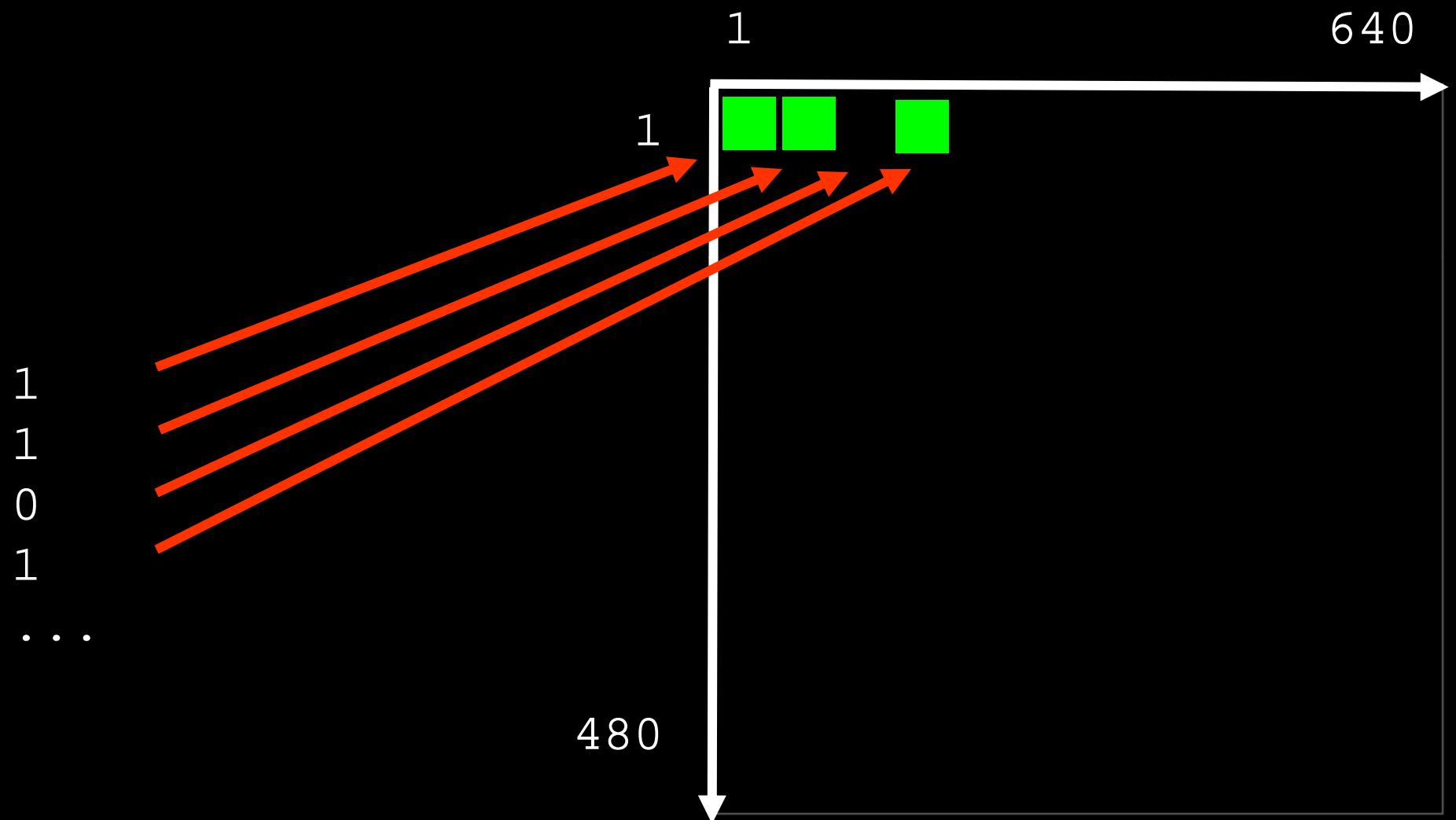


.txt

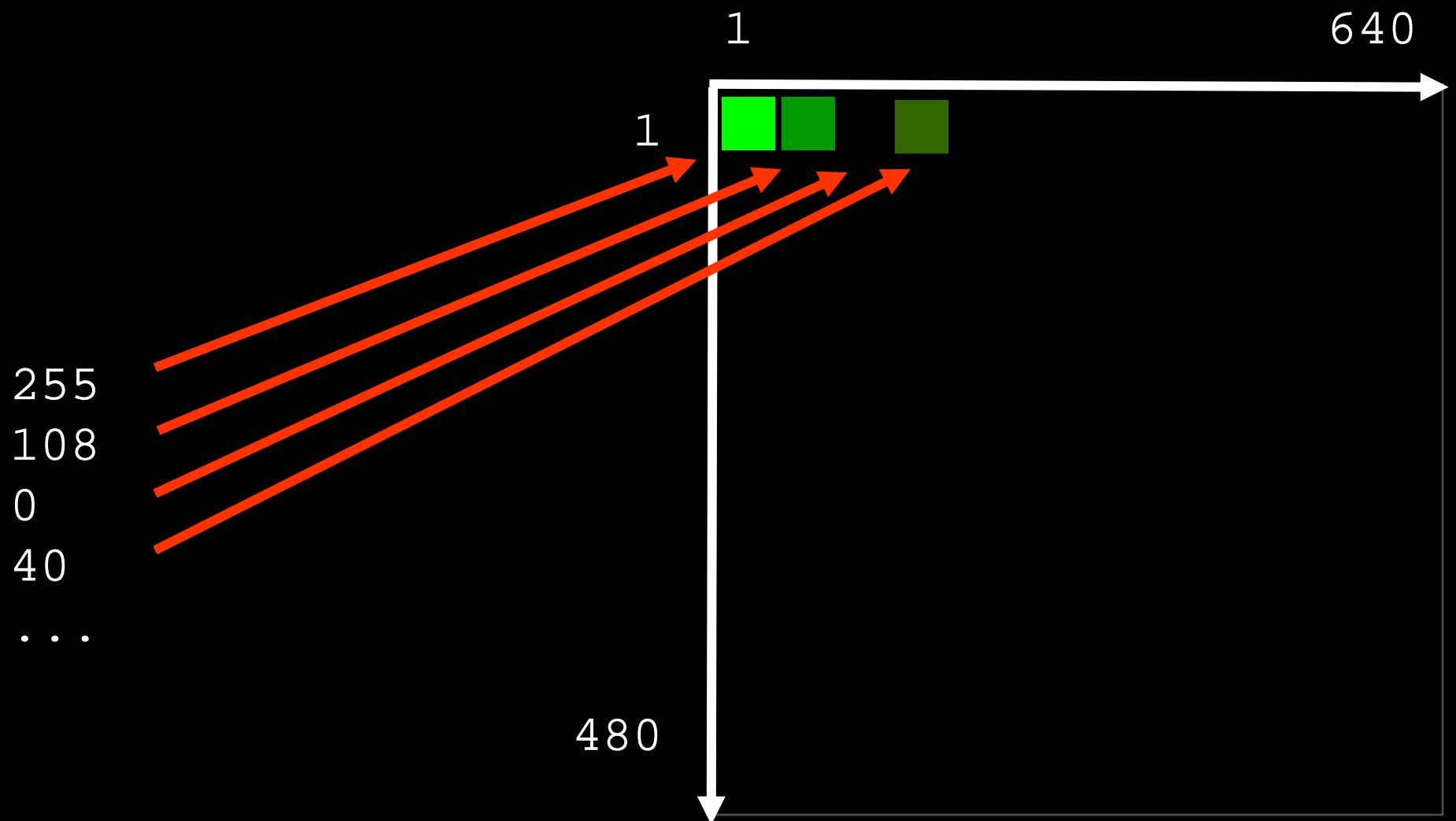


constrained pairs

Bit Plot

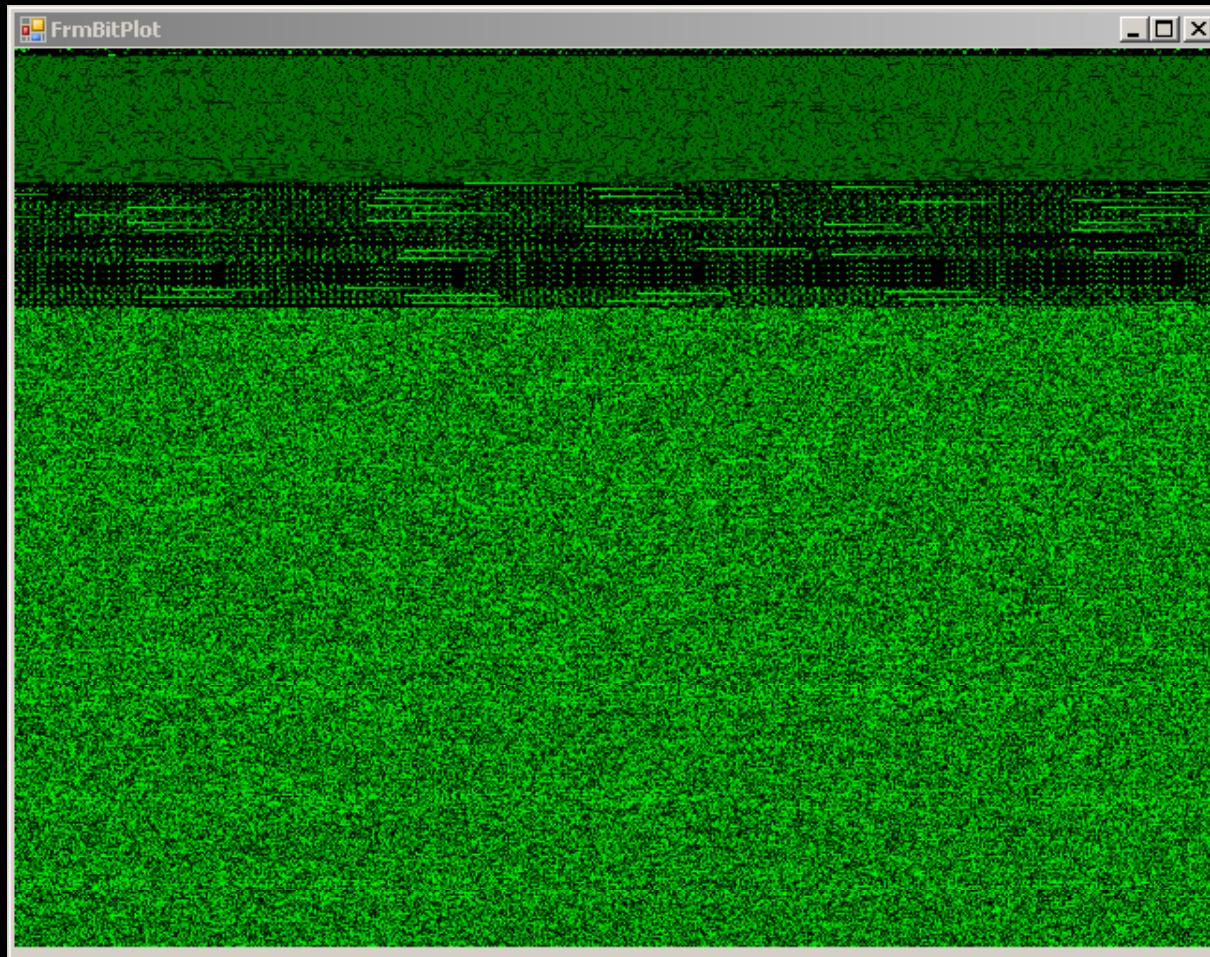


Byte Plot

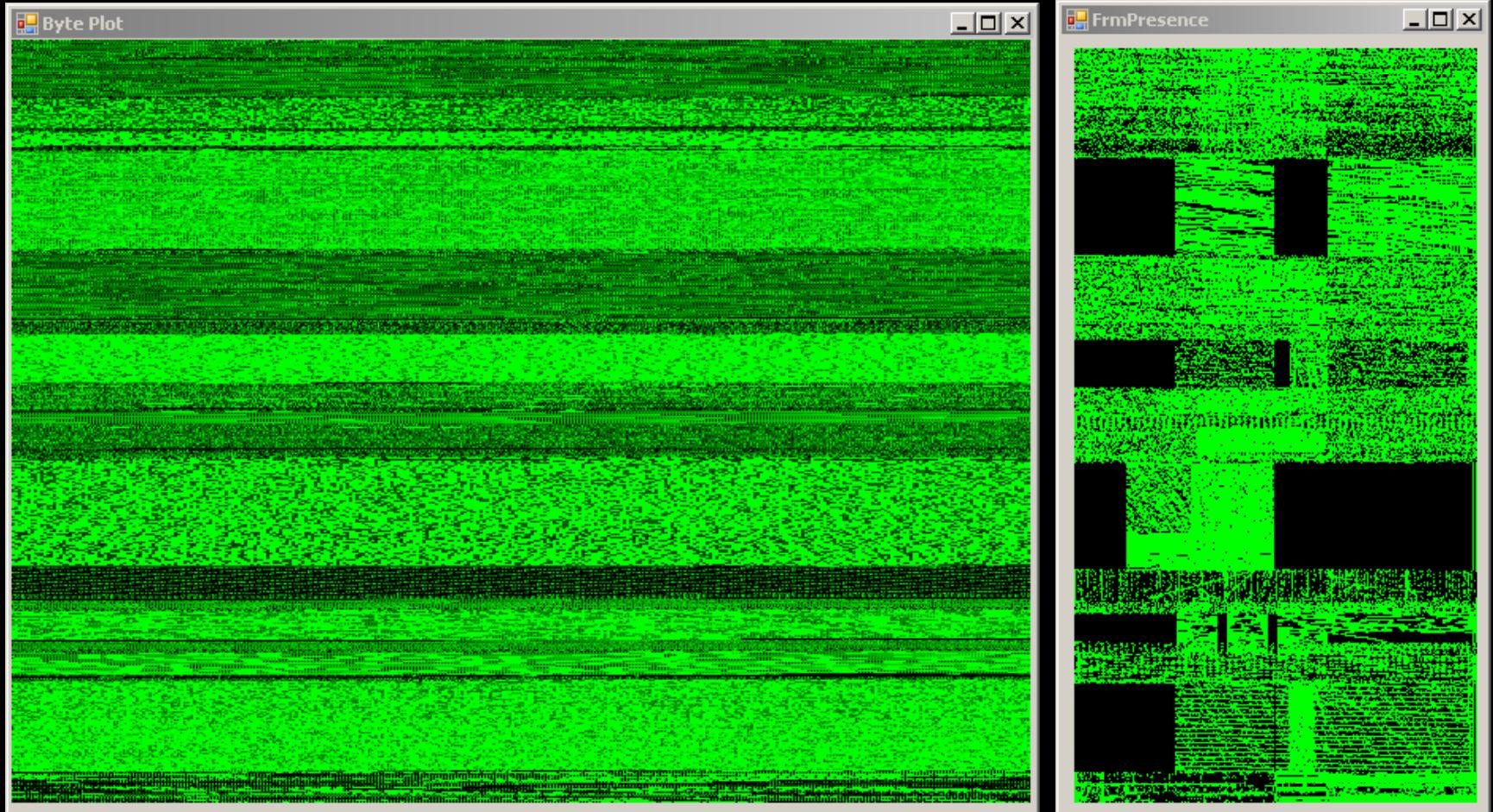


Byte Plot Example

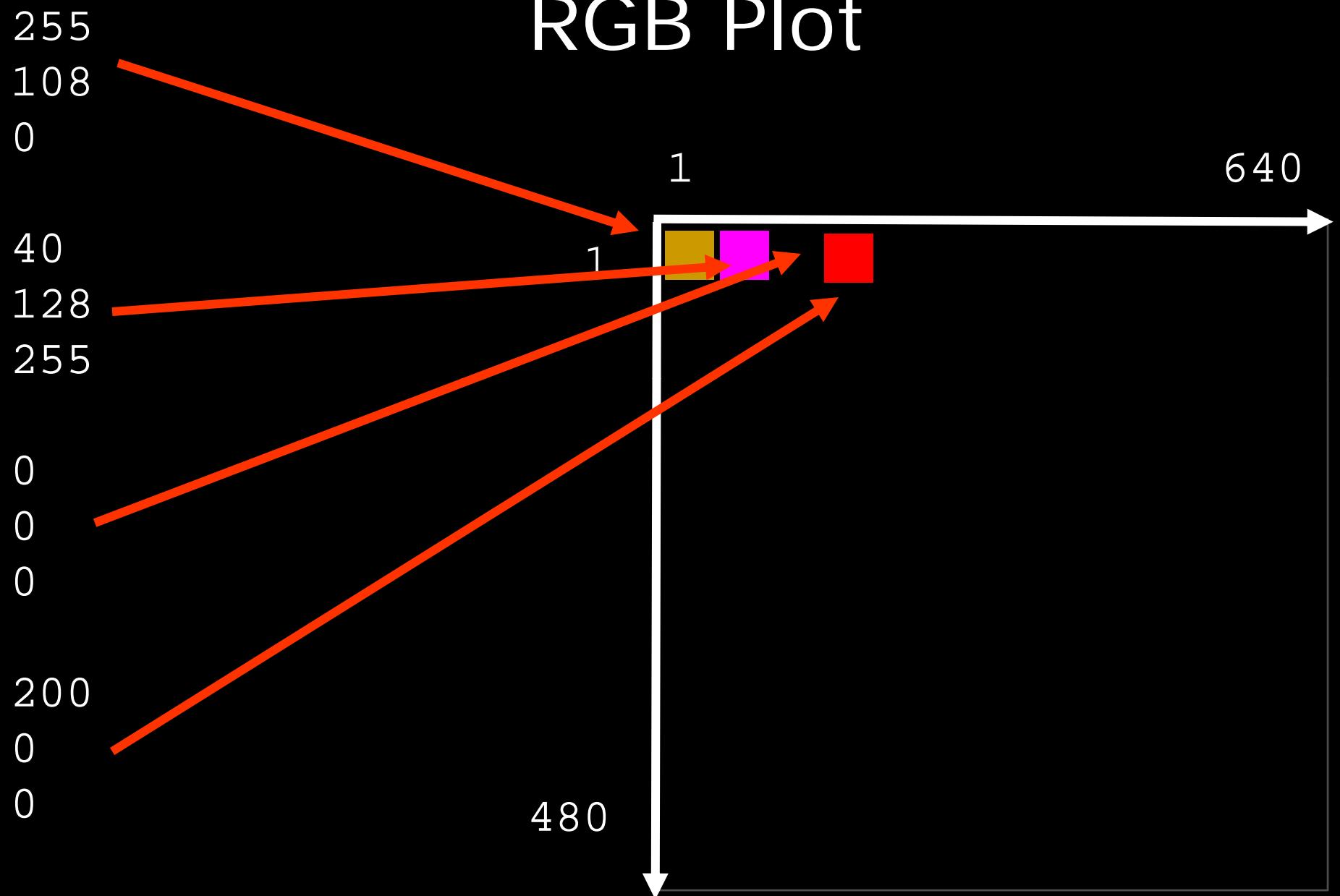
(Word Document)



Byte Presence



RGB Plot



Dot Plots

- Jonathan Helfman's "Dotplot Patterns: A Literal Look at Pattern Languages."
- Dan Kaminsky, CCC & BH 2006

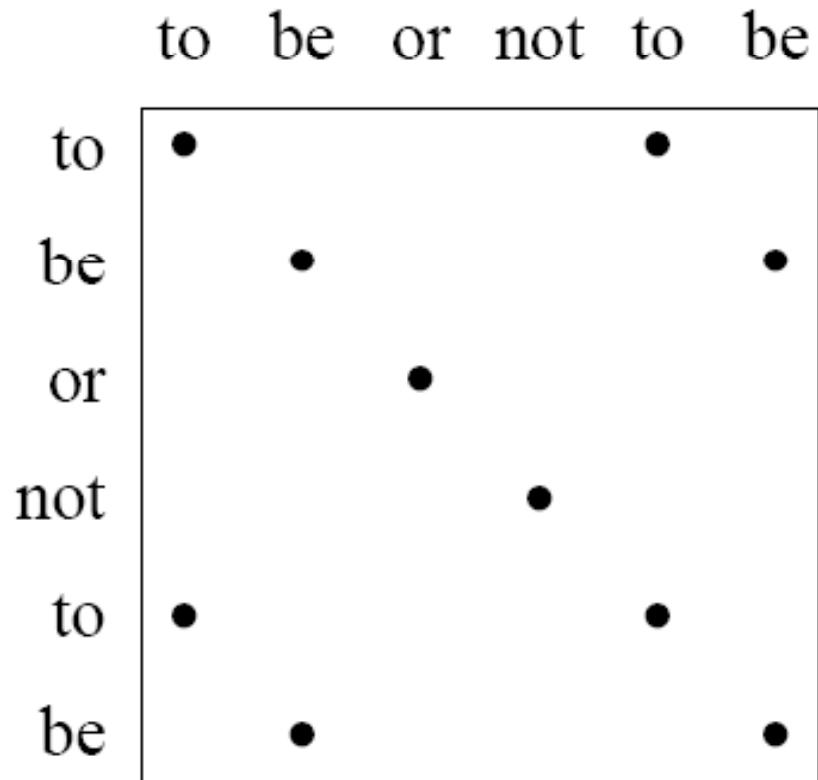
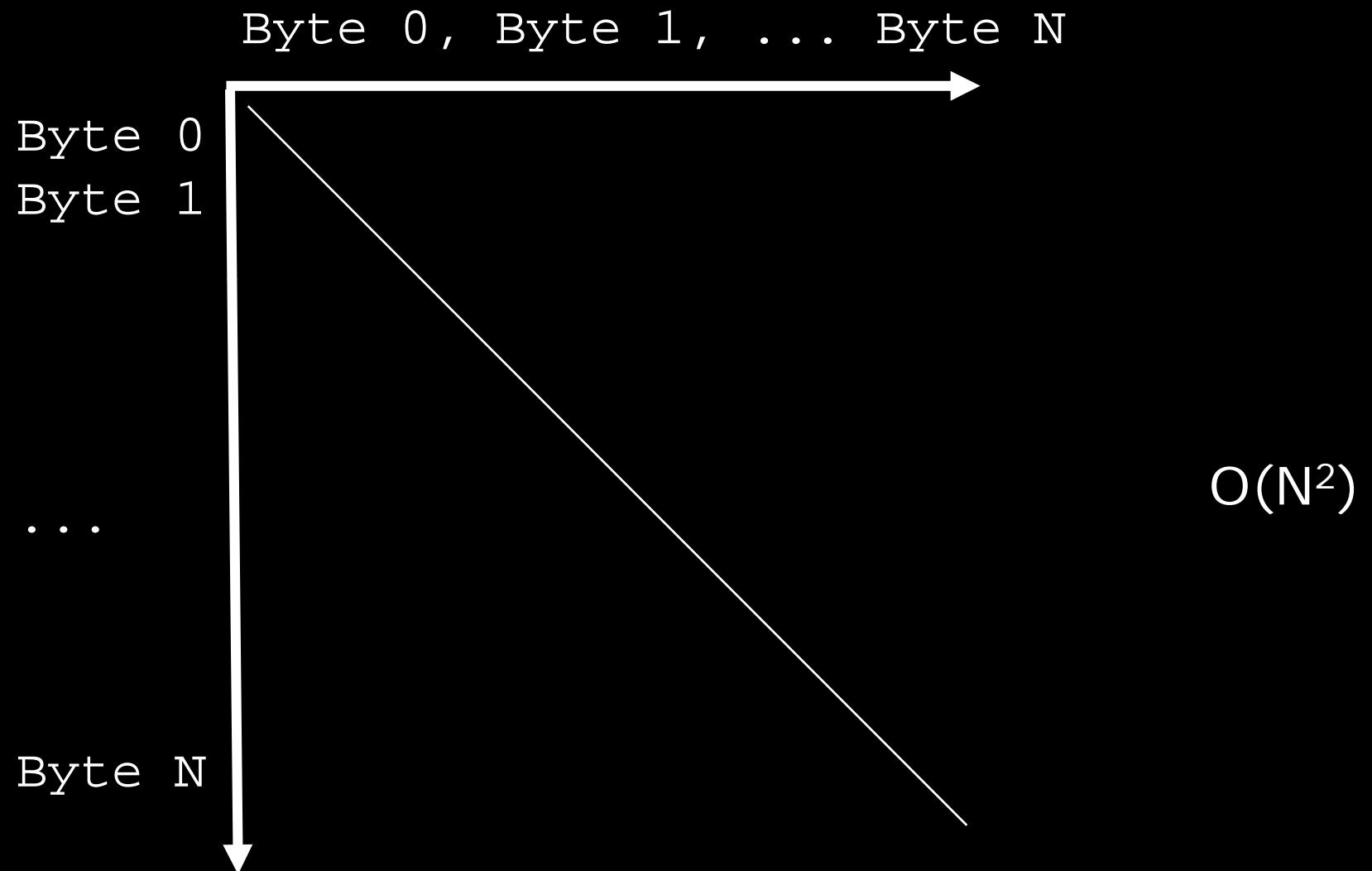
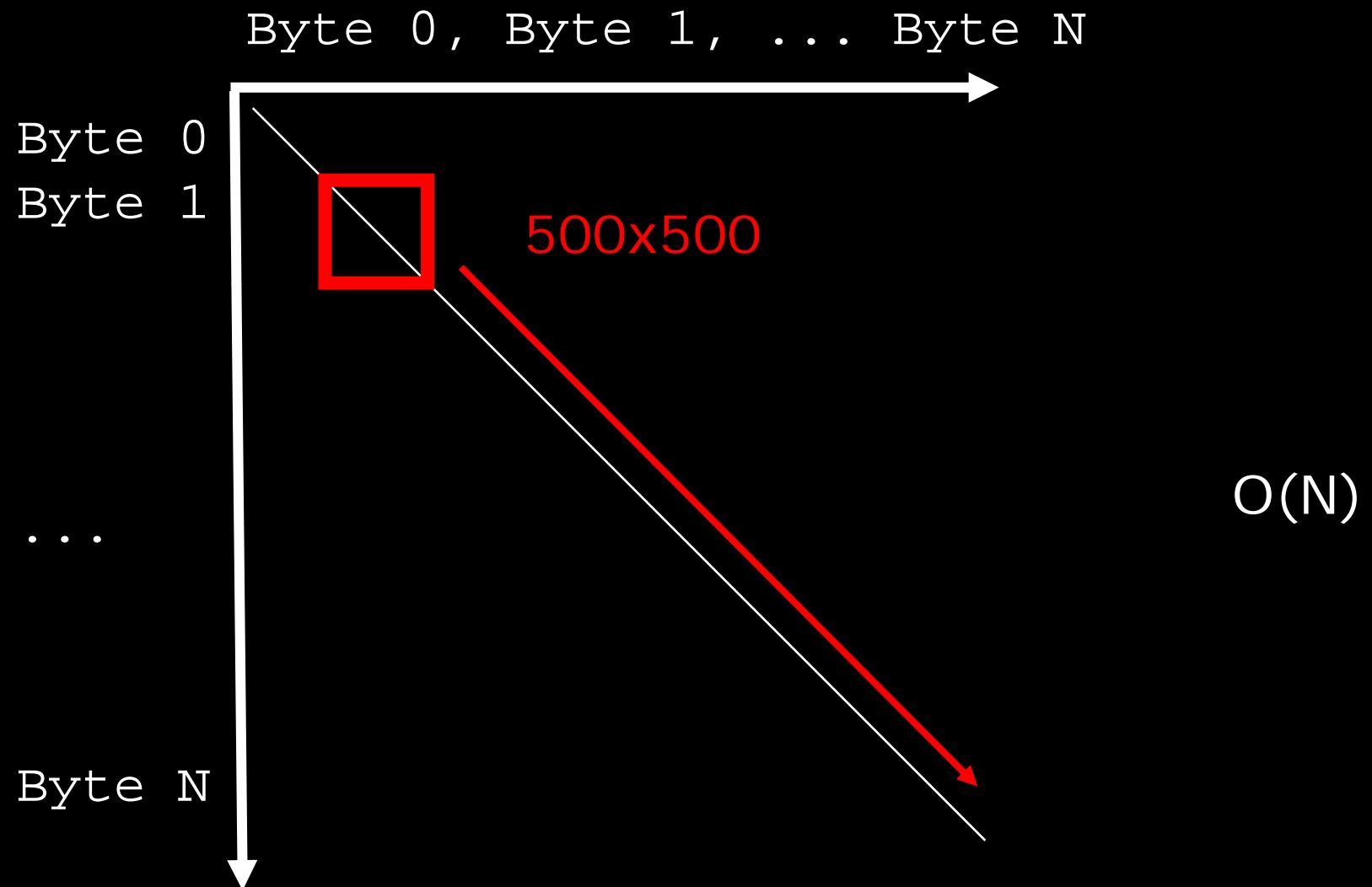


Figure 2: Six words of Shakespeare.

DotPlots

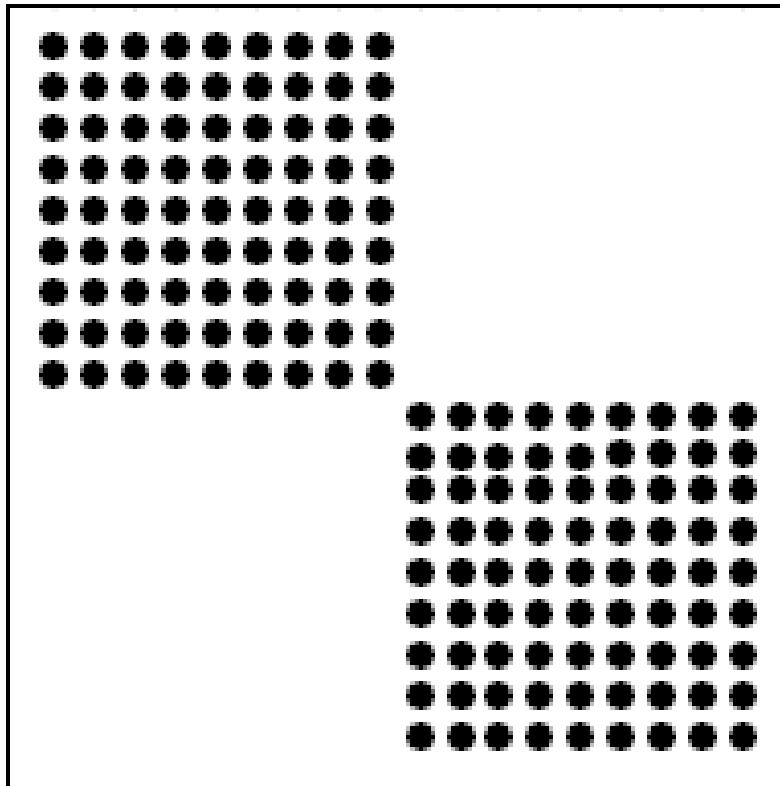


Dynamic DotPlots



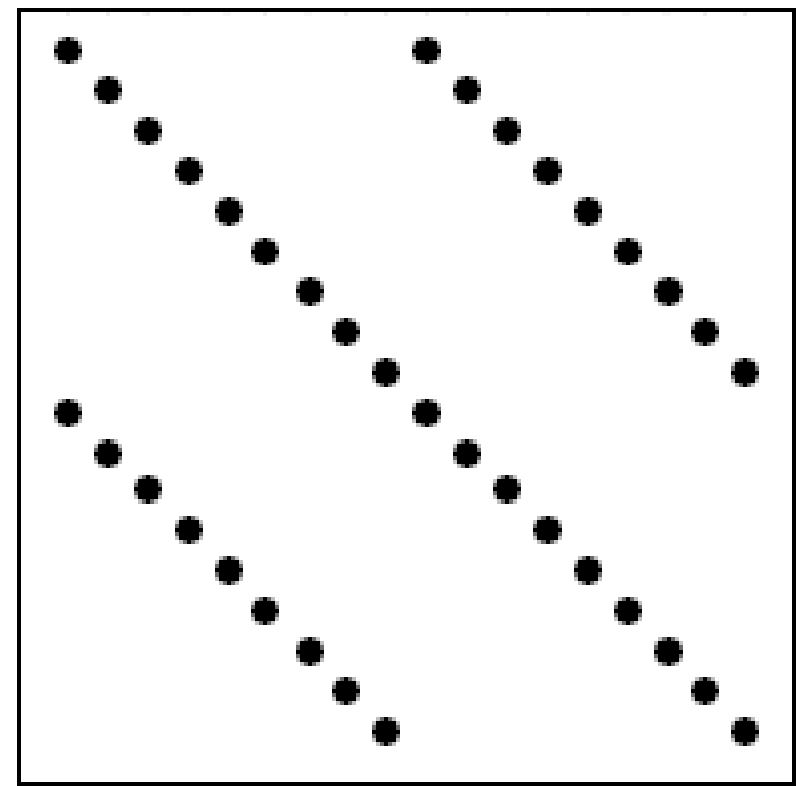
DotPlot Examples

aaaaaaaaaaaabbbbbbbbbb



a) Squares.

abcdefghijklabcdefghijkl

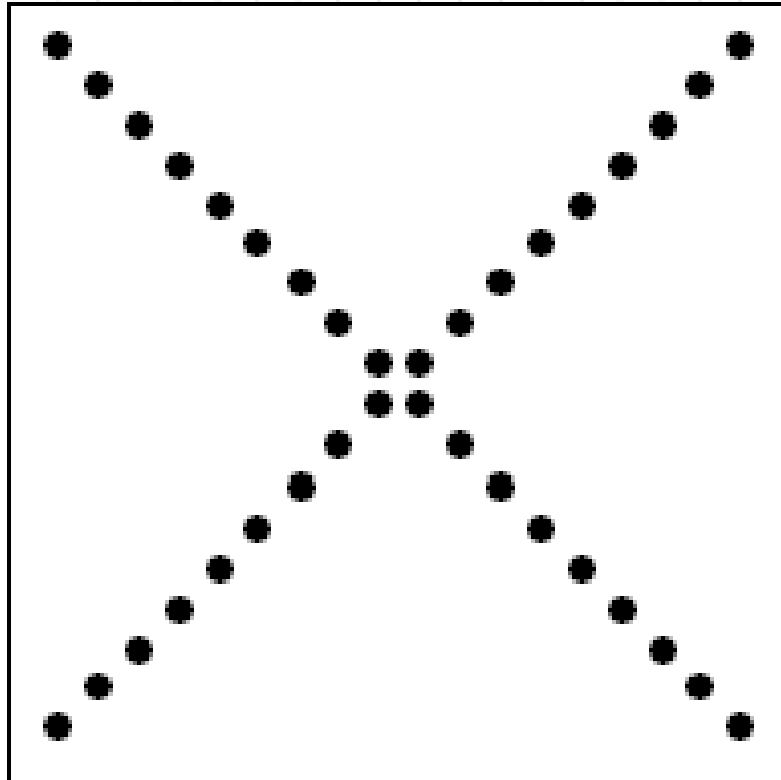


b) Diagonals.

Images: Jonathan Helfman, “Dotplot Patterns: A Literal Look at Pattern Languages.”

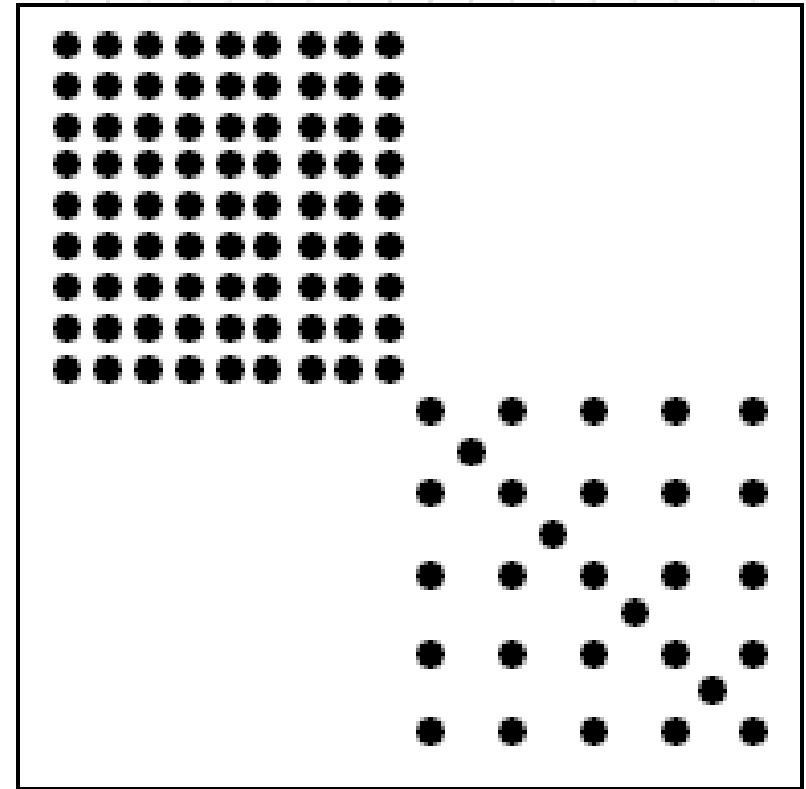
DotPlot Examples

abcdefghijklm
ihgfedcba



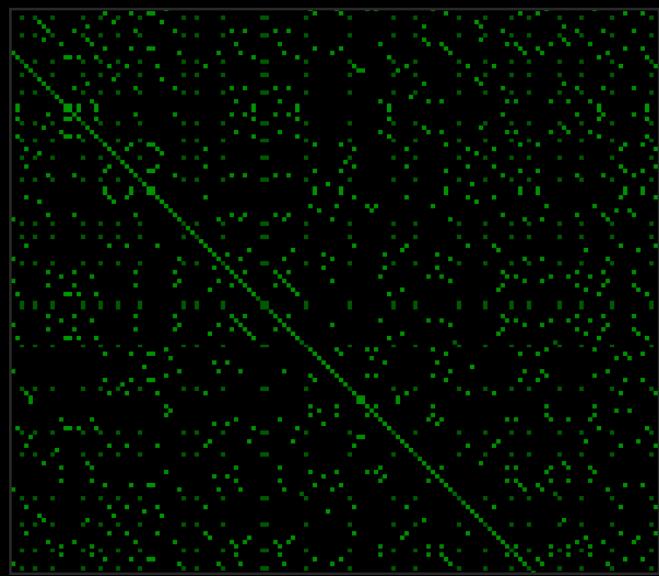
j) Palindrome.

aaaaaaaaabZbYbXbWb

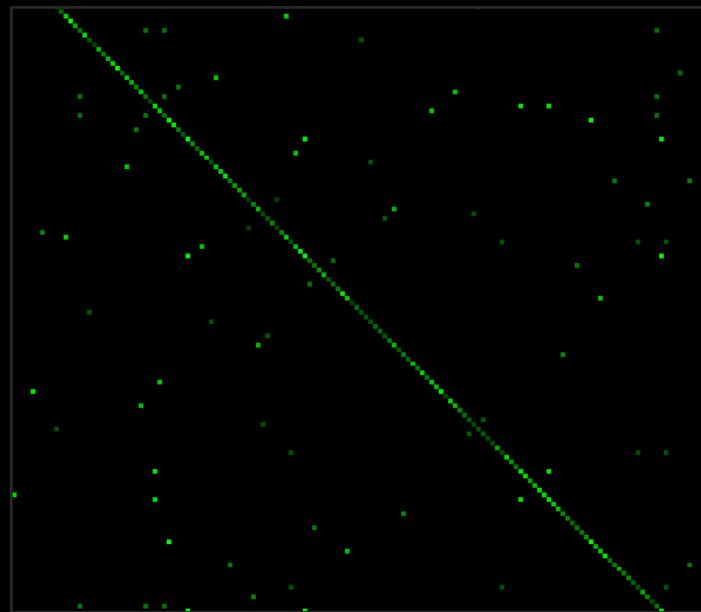


h) Density Variation.

Images: Jonathan Helfman, “Dotplot Patterns: A Literal Look at Pattern Languages.”



English Text



Compressed Audio



Bitmap Image

Byte Clouds

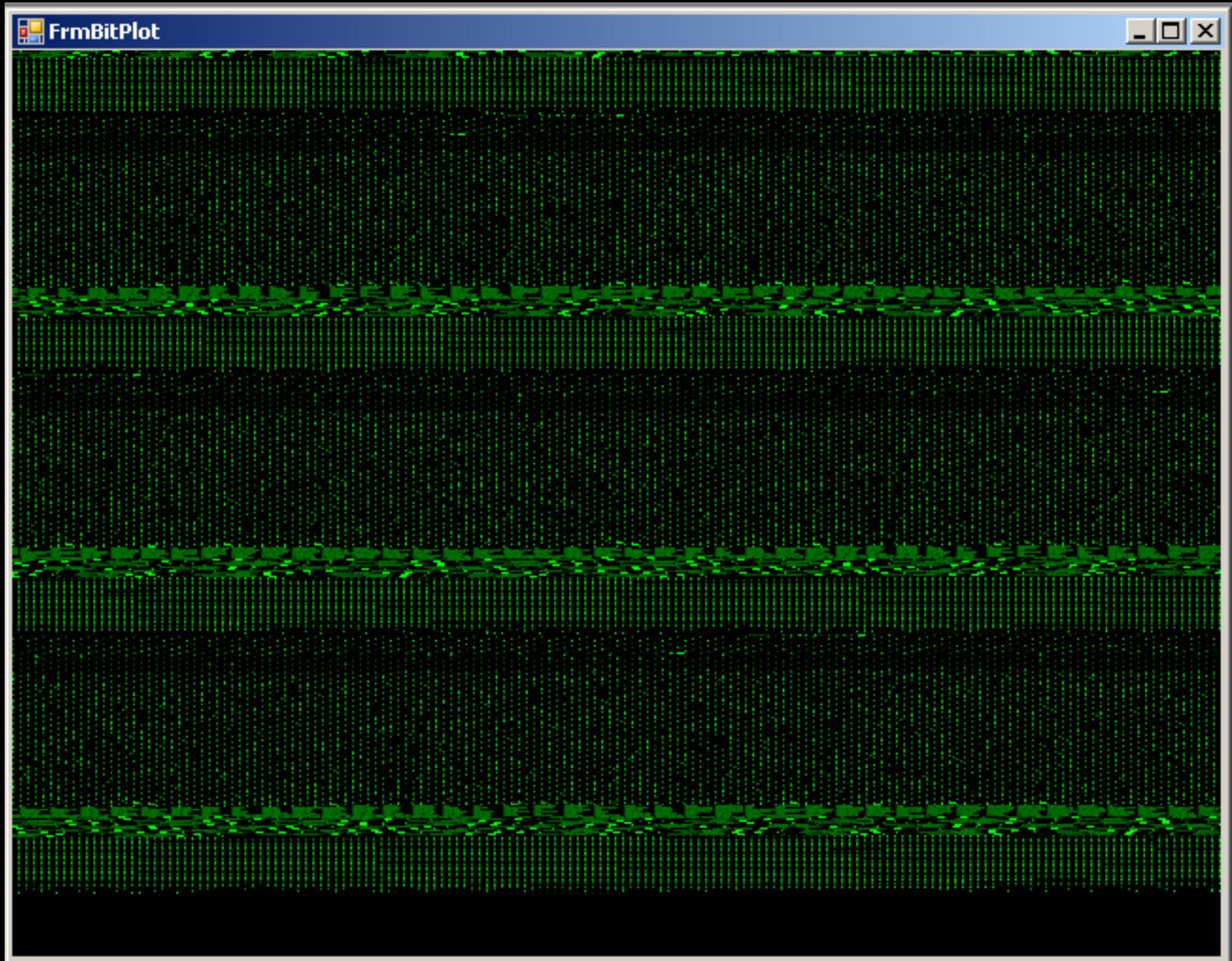
A tag cloud where words are represented by varying sizes of blue text. The most prominent words include 'addr', 'address', 'buffer', 'bytes', 'eax', 'ebp', 'ebx', 'esp', 'execve', 'int', 'main', 'movl', 'nop', 'null', 'offset', 'overflow', 'ptr', 'pushl', 'ret', 'segmentation', 'sh', 'shellcode', 'size', 'stack', 'string', 'void', and assembly-like addresses like 'x00', 'x08', 'x0b', 'x0c', 'x31', 'x46', 'x80', 'x89', 'x8d', 'xcd', and 'xff'. Smaller words include 'argc', 'argv', 'assembler', 'atoi', 'bin', 'bottom', 'bsize', 'buff', 'buffer1', 'case', 'char', 'code', 'copy', 'data', 'default', 'define', 'dump', 'edx', 'egg', 'eggsiz', 'esi', 'execute', 'function', 'gdb', 'include', 'instruction', 'jmp', 'leal', 'lets', 'looks', 'memory', 'movb', 'nopl', 'nullptr', 'offsets', 'onto', 'pointer', 'popl', 'printf', 'procedure', 'program', 'ptr', 'pushl', 'register', 'ret', 'return', 'seg', 'sh', 'shellcode', 'size', 'software', 'sp', 'start', 'strlen', 'system', 'top', 'type', 'variables', and 'vulnerable'.

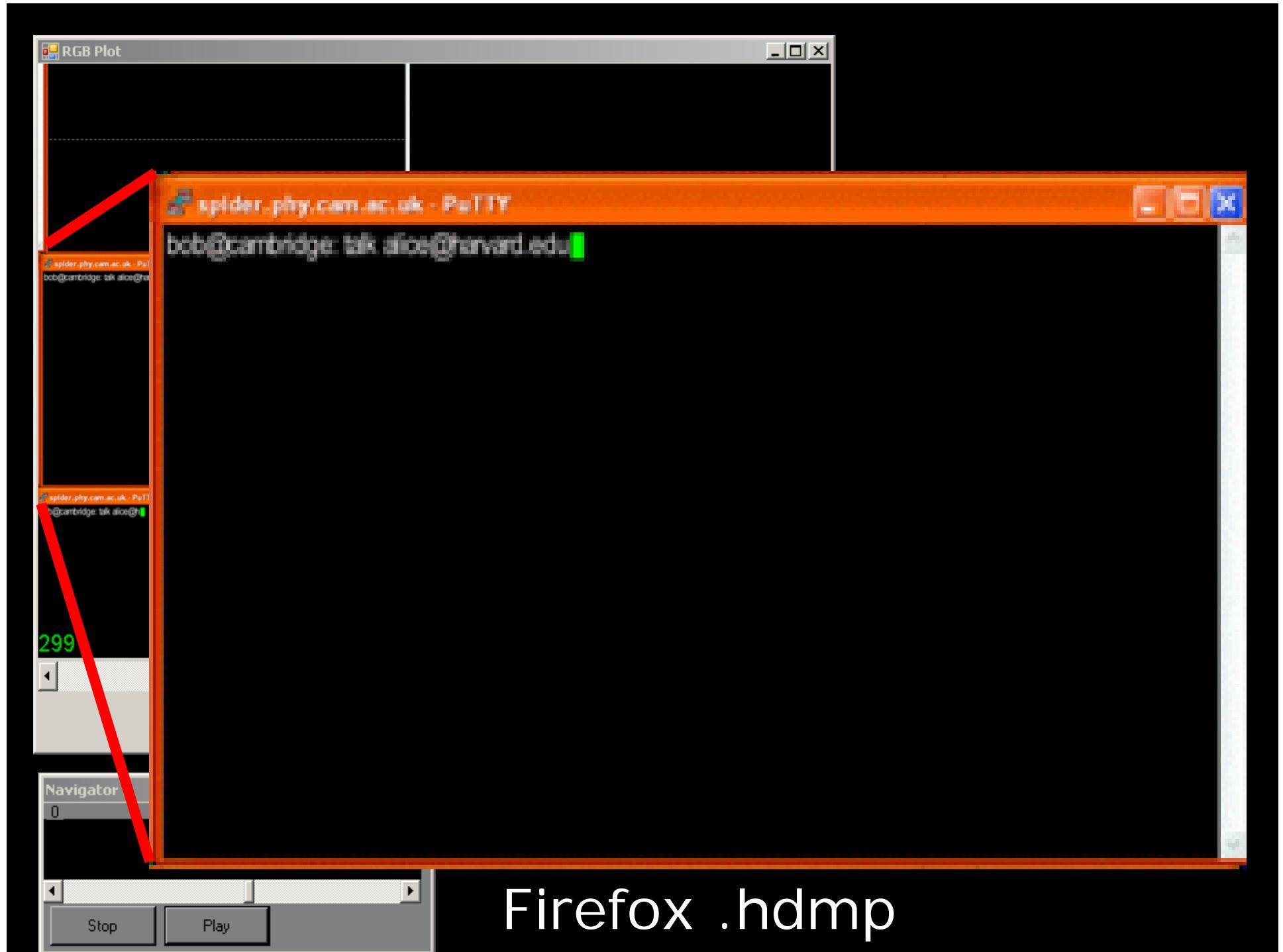
Tag Cloud
Smashing the Stack
for Fun and Profit
<http://tagcrowd.com/>



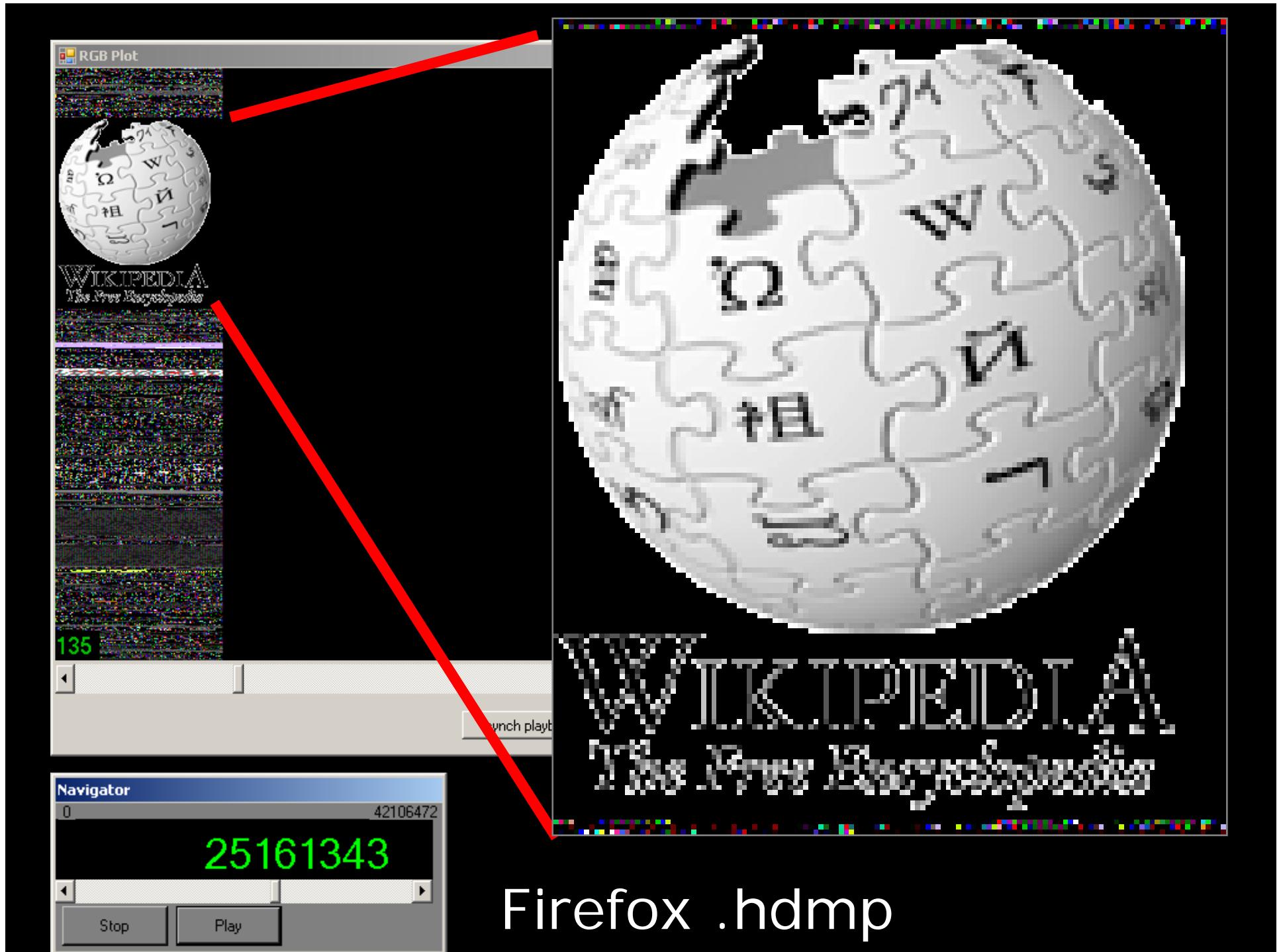
Byte Cloud

Neverwinter Nights Database File



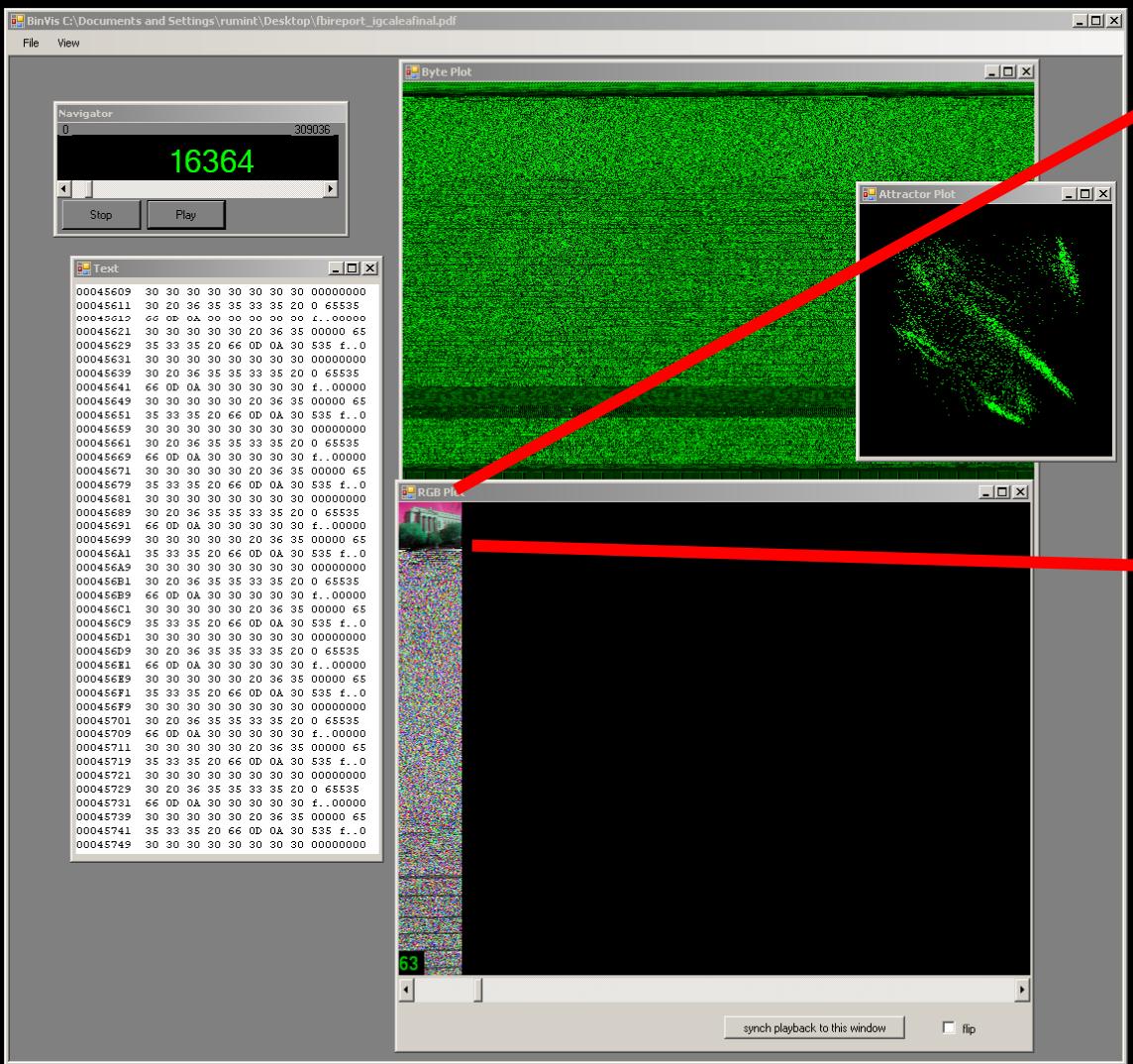


Firefox .hdmp









Redacted
PDF...

Weaknesses

- entire file may be extracted from bit/byte/RGB
 - May trigger AV or IDS
 - 8bit/byte steg
- Screams for big monitor

Demos

A Look to the Future...

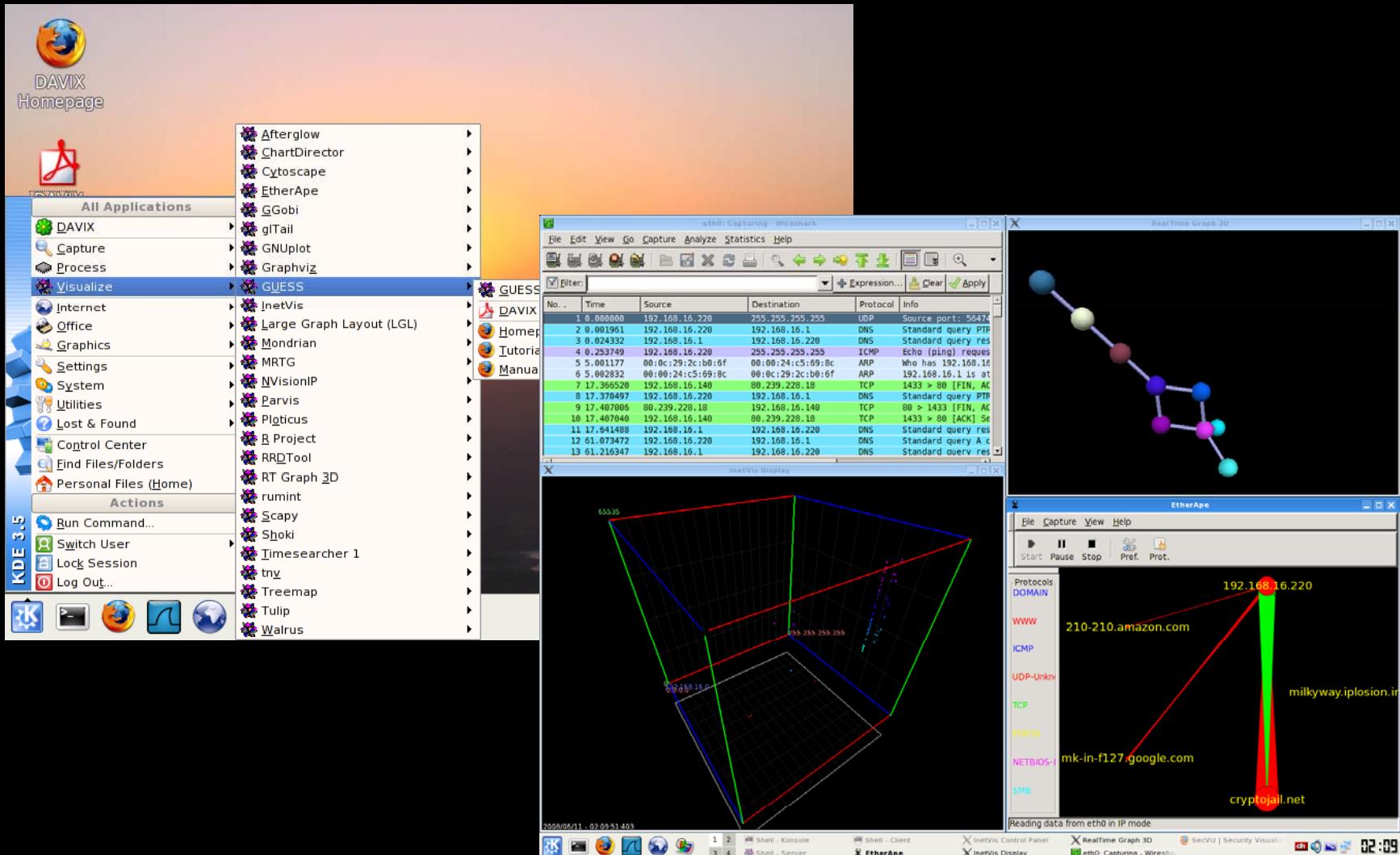
- Visual Front Ends for Offensive Tools
- Visual Cryptanalysis Support
- Human Insights Passed to Machine Processors
- User-centric Evaluation
- More Inspiration from General InfoVis Community
- Visual Fingerprints / Smart Books
- Web-based Visualization (AJAX)
- User-task Analyses
 - True Use Case Based Designs
 - Engagement of Users Beyond Students
- Examination of Full Range of Security Data
 - Merging Multiple Security Dataflows

Future Work

- Plug-ins / Editable Config Files
 - Visualizations
 - Encodings
- Saving state
 - Memory Maps
- Improving Interaction
 - What works / What doesn't
- Multiple Files / File Systems
- REGEX search
- Automated Memory Map Generation

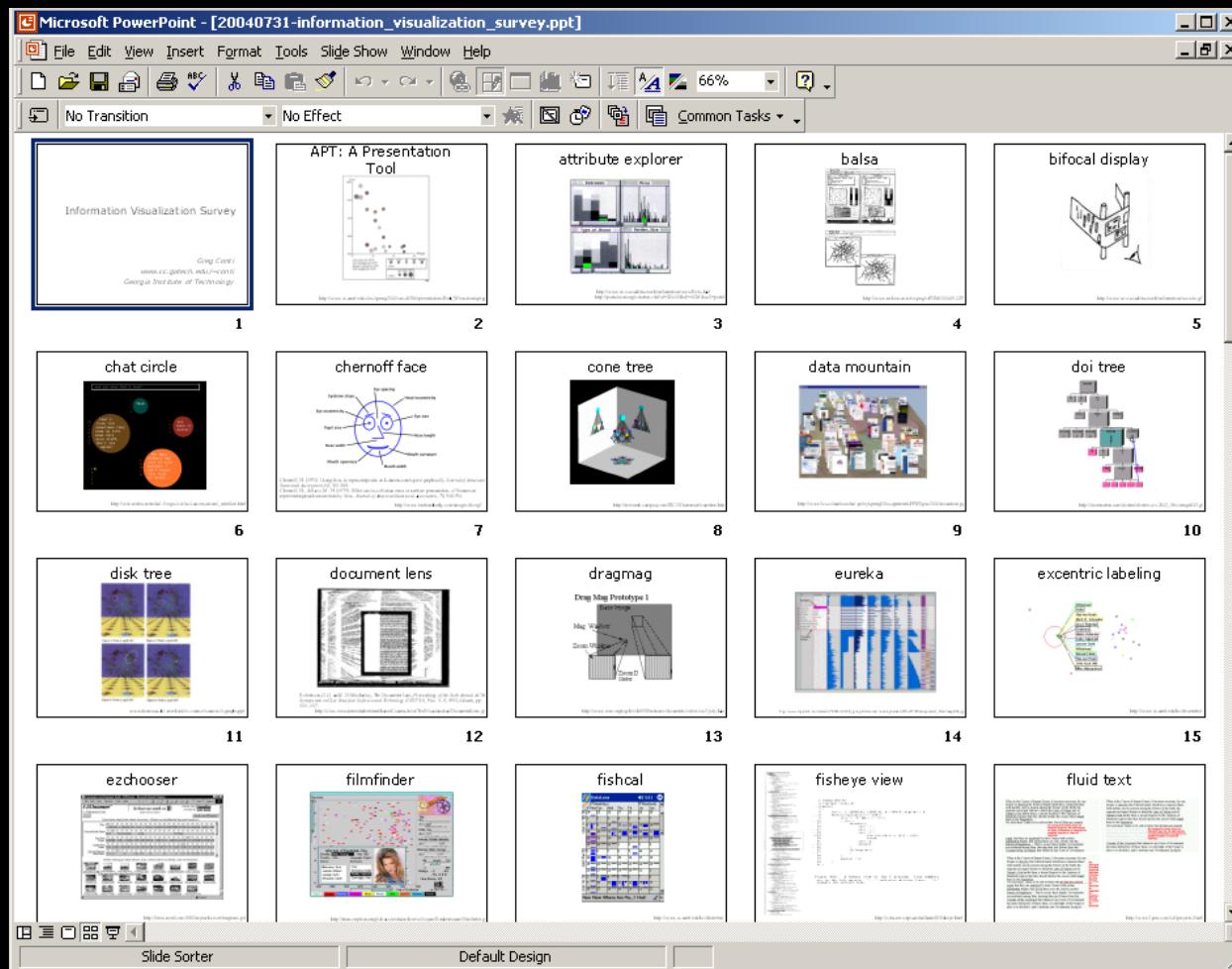
DAVIX

(Jan Monsch and Raffy Marty)

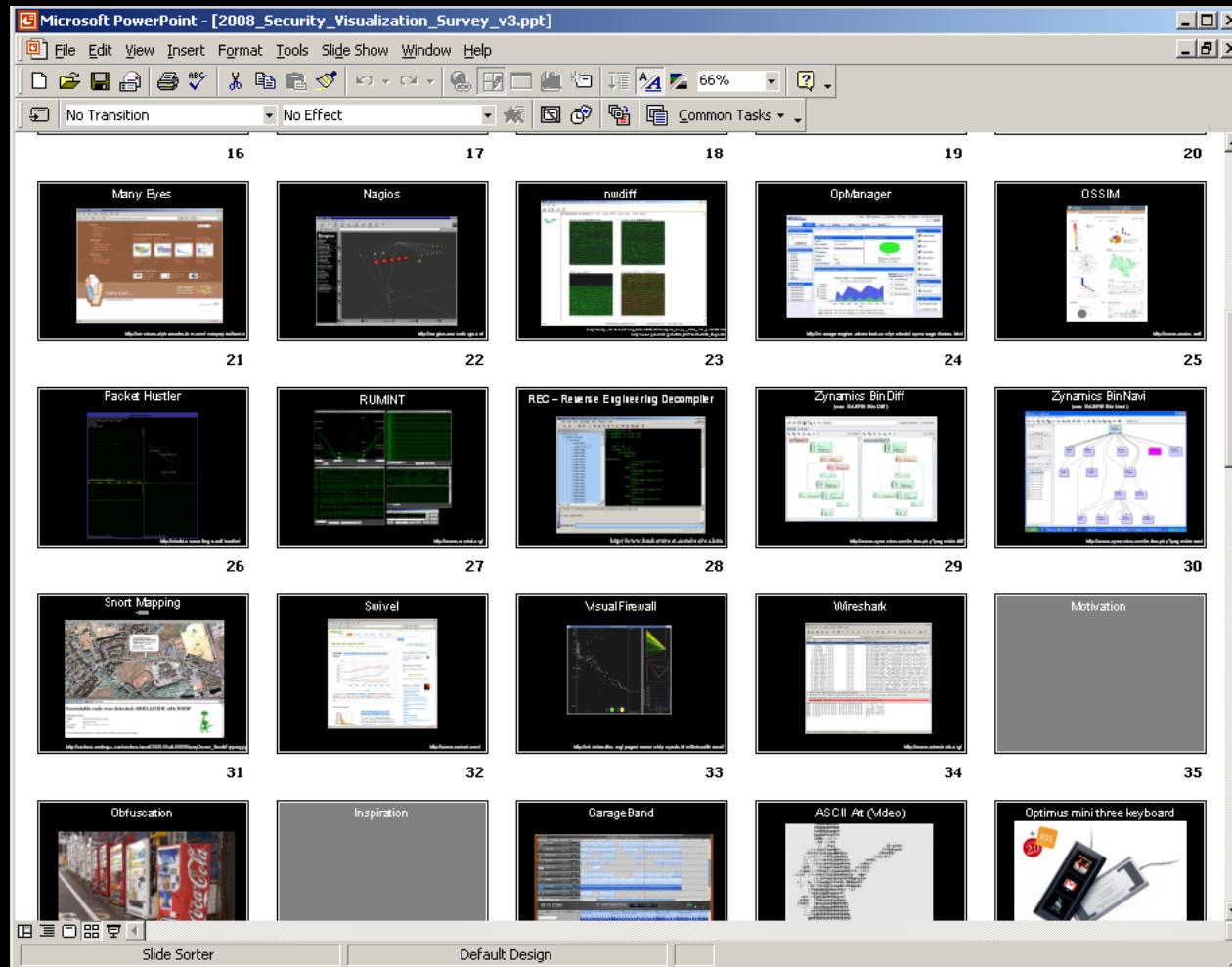


<http://www.secviz.org/node/89>

InfoVis Survey

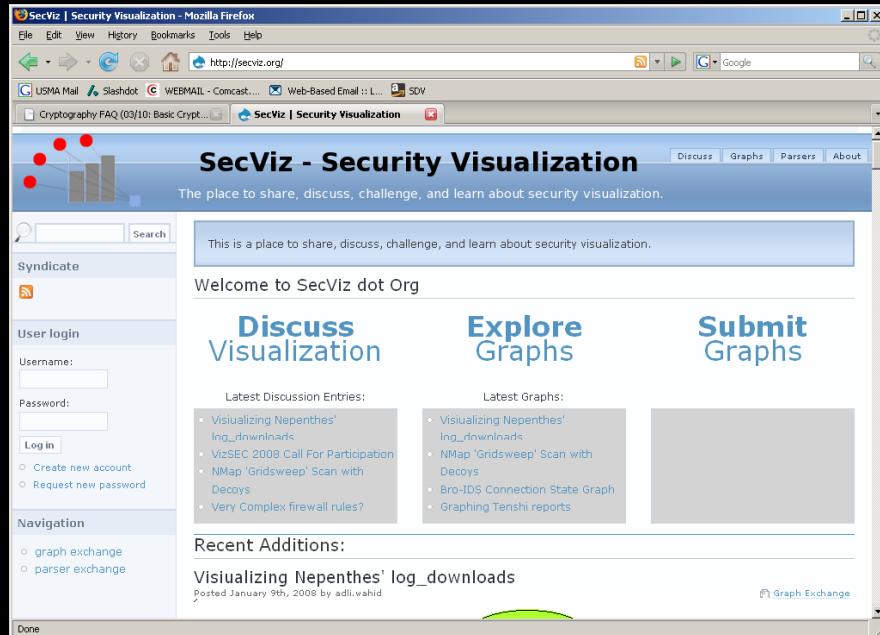


Security Visualization Survey



Communities

<http://secviz.org/>



This is a place to share, discuss, challenge, and learn about security visualization.

Welcome to SecViz dot Org

Discuss Visualization

Latest Discussion Entries:

- Visualizing Nepenthes' log_downloads
- VizSEC 2008 Call For Participation
- NMap 'Gridsweep' Scan with Decoys
- Very Complex firewall rules?

Explore Graphs

Latest Graphs:

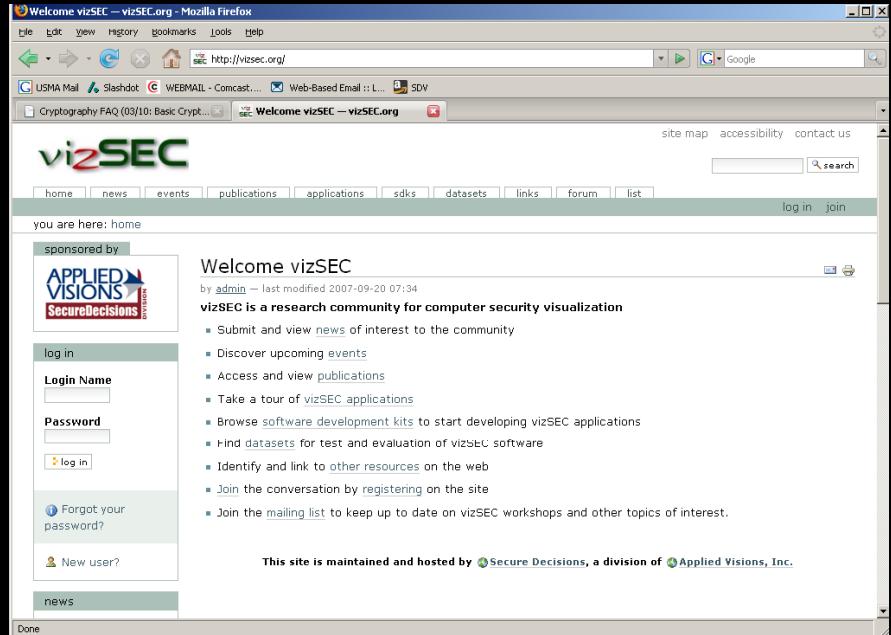
- Visualizing Nepenthes' log_downloads
- NMap 'Gridsweep' Scan with Decoys
- Bro-IDS Connection State Graph
- Graphing Tensi reports

Submit Graphs

Recent Additions:

Visualizing Nepenthes' log_downloads
Posted January 9th, 2008 by adli.wahid

<http://vizsec.org/>



The place to share, discuss, challenge, and learn about security visualization.

Welcome to SecViz dot Org

Discuss Visualization

Latest Discussion Entries:

- Visualizing Nepenthes' log_downloads
- VizSEC 2008 Call For Participation
- NMap 'Gridsweep' Scan with Decoys
- Very Complex firewall rules?

Explore Graphs

Latest Graphs:

- Visualizing Nepenthes' log_downloads
- NMap 'Gridsweep' Scan with Decoys
- Bro-IDS Connection State Graph
- Graphing Tensi reports

Submit Graphs

Recent Additions:

Visualizing Nepenthes' log_downloads
Posted January 9th, 2008 by adli.wahid

vizSEC

you are here: home

sponsored by

APPLIED VISIONS
SecureDecisions

Welcome vizSEC

by admin — last modified 2007-09-20 07:34

vizSEC is a research community for computer security visualization

- Submit and view news of interest to the community
- Discover upcoming events
- Access and view publications
- Take a tour of vizSEC applications
- Browse software development kits to start developing vizSEC applications
- Find datasets for test and evaluation of vizSEC software
- Identify and link to other resources on the web
- Join the conversation by registering on the site
- Join the mailing list to keep up to date on vizSEC workshops and other topics of interest.

This site is maintained and hosted by Secure Decisions, a division of Applied Visions, Inc.

“The place to share, discuss, challenge, and learn about security visualization.”

Raffy Marty
Splunk

“vizSEC is a research community for computer security visualization.”

John Goodall
Secure Decisions

VizSEC 2008

VizSEC 2008 Workshop on Visualization for Cyber Security - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.vizsec.org/workshop2008/ Google

USMA Mail Slashdot WEBMAIL - Comcast... Web-Based Email :: L... SDW

Cryptography FAQ (03/10: Basic Crypt... VizSEC 2008 Workshop on Visuali...)

VizSEC 2008

Workshop on Visualization for Computer Security



September 15, 2008 / Cambridge, MA USA
In conjunction with [RAID 2008](#)

Call for Papers

The 5th International Workshop on Visualization for Cyber Security will provide a forum for new research in visualization for computer security. We are pleased to be holding this year's meeting in conjunction with [the 11th International Symposium on Recent Advances in Intrusion Detection](#). The VizSEC Workshop will be held at MIT in Cambridge, Massachusetts USA on Monday, September 15, 2008.

As a result of previous VizSEC workshops, we have seen both the application of existing visualization techniques to security problems and the development of novel security visualization approaches. However, VizSEC research has focused on helping *human analysts* to detect anomalies and patterns, particularly in

Done

<http://www.vizsec.org/workshop2008/>

More Information

- “Visual Reverse Engineering of Binary and Data Files.” Gregory Conti, Erik Dean, Matthew Sinda, Benjamin Sangster. VizSEC 2008.
 - Available September
- Security Data Visualization (No Starch Press)
- Applied Security Visualization (Addison-Wesley)

Visual Reverse Engin

Gregory Conti, Erik Dean

Department of Electric
United States

We

{gregory.conti, erik.dean, m...

Abstract.

The analysis of computer files performed by security professionals seeking to detect and analyze malicious code and other harmful formats for their products, and for law enforcement agencies investigating computer crimes. The behavior and structure of undocumented software is often analyzed using tools such as disassemblers, editors, and debuggers. These tools are typically used in conjunction with text based approaches. In this paper we present a visual approach which support meaningful investigation of binary and data files.

Acknowledgements

Damon Becknell, Jon Bentley, Jean Blair, Sergey Bratus, Chris Compton, Tom Cross, Ron Dodge, Carrie Gates, Chris Gates, Joe Grand, Julian Grizzard, Toby Kohlenberg, Oleg Kolesnikov, Frank Mabry, Raffy Marty, Brent Nolan, Gene Ressler, Ben Sangster, Matt Sinda, and Ed Sobiesk

"In fact, master reversers like Fravia recommend cracking while intoxicated with a mixture of strong alcoholic beverages.

While for health reasons we cannot recommend this method, you may find that a relaxing cup of hot tea unwinds your mind and allows you to think in reverse."

-from *Security Warrior*