Information Visualization for Cyber Security

John Goodall
Oak Ridge National Lab
Computational Sciences and Engineering Division
The Good…

Average household

- Mortgage $84,911
- Auto and tuition loans $14,414
- Credit cards $8,565
- Home equity loans $10,062

Savings each year

- $392

Source: Federal Reserve

The Bad...

...and if the decline was fast, the recovery took a considerable time. Most bear markets hit bottom after a decline of three years or less... and it usually takes less than 10 years to rebound to the previous high.

MARKET PEAK BEFORE MELTDOWN

If you consider March 9 to be the bottom of the present meltdown, that would mean the slide lasted only 17 months, a short period compared to the average.

It took nearly three years for the market to reach the bottom after the 1929 crash... and 22 years to bounce back to its original level.

It took four years to recover from the 2000 meltdown.

The Dow gained 19% since March 9.

MONTHS THE MARKET TOOK TO REACH BOTTOM
YEARS THE MARKET TOOK TO RETURN TO INITIAL LEVEL

SOURCES: Dow Jones indexes
And The Ugly.
London, 1854
Visualization can convince your strongest opponents.
Russia, 1812
Carte Figurative des pertes successives en hommes de l'Armée Francaise dans la campagne de Russie 1812 - 1813.

Paris, le 20 Novembre 1869.

Les nombres d'hommes perdu sont représentés par les longueurs des zénes colorées à raison de 1'000'000 pour 500 hommes ; ils sont le plus locaux en suivant des zénes. Le temps, pendant les heures qui suivraient le retour de l'Armée, le mieux et en situation. Les mouvements qui se sont vus à travers le carte ne se produisent qu'aux environs de Moscou, de Kiev, de Tchernigov, de Charkov, et le journal suivi de l'Armée, jusqu'au 28 Octobre.

Les mois sont finis, jour à l'end de l'état de l'Armée, j'ai suivi que les corps de Prince Napoléon et du Maréchal Davoust, qui avaient été attachés sur Moscou et son environs, nous avons toujours marchés avec l'Armée.
Visualization can tell a story
Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information.

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.

The communication of abstract data through the use of interactive visual interfaces.

Information visualization utilizes computer graphics and interaction to assist humans in solving problems.

Information visualization promises to help us speed our understanding and action in a world of increasing information volumes.

Visual representations of the semantics, or meaning, of information. In contrast to scientific visualization, information visualization typically deals with nonnumeric, nonspatial, and high-dimensional data.

The purpose of visualization is insight, not pictures.
Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information.

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.

The communication of abstract data through the use of interactive visual interfaces.

Information visualization utilizes computer graphics and interaction to assist humans in solving problems.

Information visualization promises to help us speed our understanding and action in a world of increasing information volumes.

Visual representations of the semantics, or meaning, of information. In contrast to scientific visualization, information visualization typically deals with nonnumeric, nonspatial, and high-dimensional data.

The purpose of visualization is insight, not pictures
Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information.

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.

The communication of abstract data through the use of interactive visual interfaces.

Information visualization utilizes computer graphics and interaction to assist humans in solving problems.

Information visualization promises to help us speed our understanding and action in a world of increasing information volumes.

Visual representations of the semantics, or meaning, of information. In contrast to scientific visualization, information visualization typically deals with nonnumeric, nonspatial, and high-dimensional data.

The purpose of visualization is insight, not pictures.
Information visualization is a set of technologies that use visual computing to amplify human cognition with abstract information.

The use of computer-supported, interactive, visual representations of abstract data to amplify cognition.

The communication of abstract data through the use of interactive visual interfaces.

Information visualization utilizes computer graphics and interaction to assist humans in solving problems.

Information visualization promises to help us speed our understanding and action in a world of increasing information volumes.

Visual representations of the semantics, or meaning, of information. In contrast to scientific visualization, information visualization typically deals with nonnumeric, nonspatial, and high-dimensional data.

The purpose of visualization is insight, not pictures.
Color

Find red circle
Find red circle
Find red circle
Color

Find red circle
Shape

Find red circle
Shape

Find red circle
Shape

Find red circle
Find red circle
Conjunction: Color and Shape

Find red circle
Find red circle
Conjunction: Color and Shape

Find red circle
Conjunction: Color and Shape

Find red circle
$stocks by Sector
http://www.smartmoney.com/map-of-the-market/
Visualization can enable insights.
Cyber Security
Cyber now is part of our social fabric

Largest data breaches, 2000-2009

State sponsored cyber attacks

Moonlight Maze (Russia, 1999)
Titan Rain (China, 2003-?)
GhostNet (China, ?-2009)

Estonian DDoS Attacks (Russia, 2007)
Russia–Georgia War (Russia, 2008)
Oplan 5027 (N. Korea, 2009)
network monitoring intrusion detection trend analysis binary code analysis correlation analysis vulnerability analysis application monitoring attack paths forensic analysis attack analysis routing analysis packet tracing host monitoring software code review malware analysis cryptography traffic classification protocol analysis log monitoring
network monitoring  intrusion detection
  trend analysis  binary code analysis
  correlation analysis  vulnerability analysis
  application monitoring  attack paths
  forensic analysis  attack analysis  routing
  analysis  packet tracing  host monitoring
  software code review  malware analysis
  cryptography  traffic classification  protocol
  analysis  log monitoring
Research Examples
Visual Analysis for Computer Network Defense
1,542,761,868*

* Number of flows captured on one day at US-CERT
What is the state of the art in network traffic analysis tools for the entire US Government?
What is the state of the art in network traffic analysis tools for the US Government?
Netflow Visual Analysis
1. Select fields to fetch
2. Select fields to filter
3. Select ranges of numerical data
4. Select categorical data
5. Relate selections
6. Select logical operations
A Visual Query Builder: Simplifying Data Selection

Steps to Build a Query

1. Select fields to fetch
2. Select fields to filter
3. Select ranges of numerical data
4. Select categorical data
5. Relate selections
6. Select logical operations

Robert Ferris
Applied Vision, Inc.
Securite Decisions Division
6 Bayview Avenue
Hauppauge, NY 11768
(631) 754-4920
RobertF@SecuriteDecisions.com

Brianne O.
Securite Decisions
6 Bayview Avenue
Hauppauge, NY 11768
(631) 759-3908 (Direct)
BrianneO@SecuriteDecisions.com

Contact
DHS Contract # FA8750-08-C-0140

Alternate Contact:
Benefits

• Simplifies query building for both non-expert and expert users
• Reduces error and guesswork in query construction
• Improves efficiency, saves time
• Allows discovery of things that were previously invisible
• Lets you see what’s there before you start querying
Visual Analysis
Benefits

• Multiple perspectives allow analysts to see data in new ways and put attacks into context
• Simple visualizations are intuitive
• Powerful filtering encourage exploration
• Big picture view reveals trends
Software Assurance Visualization
More than 98% of all PCs have one or more vulnerable programs

Software Assurance: poorly written software is at the root of all of our security problems

Doug Maughan, CACM 53(2)
Top 10 Hard Problems in Cyber Security
More than 98% of all PCs have one or more vulnerable programs

http://secunia.com/blog/56/

Lots of Bad Code

Software Assurance: poorly written software is at the root of all of our security problems

Doug Maughan, CACM 53(2)
Top 10 Hard Problems in Cyber Security
“No tool stands out as an uber-tool. Each has its strengths and weaknesses.”

Kris Britton, Technical Director
NSA’s Center for Assured Software

84% of the vulnerabilities were identified by one tool and one tool alone.
No tool stands out as an **uber-tool**. Each has its **strengths** and **weaknesses**.

---

**No Tool is Perfect**

84% of the vulnerabilities were identified by **one tool** and **one tool alone**

---

Kris Britton, Technical Director
NSA's Center for Assured Software
Tools find different vulnerabilities

- Non-overlap: Hits reported by one tool and no others (84%)
- Overlap: Hits reported by more than one tool (16%)
  - 2 tools
  - 3 tools
  - 4 tools
  - All 5 tools

*from MITRE*
Tools find different vulnerabilities

Non-overlap: Hits reported by one tool and no others (84%)

Overlap: Hits reported by more than one tool (16%)

from MITRE
Technical Approach

Software Assurance Visual Analysis

Provide a workflow for developers to
bring together disparate security analysis results
visually analyze and prioritize those results
explore those results to uncover hidden trends
use code context to assess the impact of those results
see who is responsible for vulnerabilities
assign vulnerabilities to developers responsible
Use case : Triage

- Which vulnerabilities are noise / most important
- What vulnerability categories are most common
- What vulnerabilities are found by multiple tools
- Where in the code are the vulnerabilities
- Who do confirmed vulnerabilities get assigned to
Visualization

- Each source code file is represented as a block
- Each block aggregates the vulnerabilities found
- Very compact, space filling method
- Flexible (color/sort) data > visual mappings
Overview first, zoom & filter, details on demand...

– Ben Schneiderman
### StandardWrapperValve.java

**Location:** container\catalina\src\share\org\apache\catalina\core

- **Severity Score:** 8.25
- **Minimum Severity:** 6
- **Maximum Severity:** 10

#### Vulnerabilities

<table>
<thead>
<tr>
<th>Objects</th>
<th>Severity Score</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>StandardWrapperValve</code></td>
<td>8.25</td>
<td></td>
</tr>
<tr>
<td>public invoke(Request request, Response response, ValveContext val)</td>
<td>8.25</td>
<td></td>
</tr>
<tr>
<td>Line 192 - Errors: NPE_COND (Klocwork)</td>
<td>10.00</td>
<td></td>
</tr>
<tr>
<td>Line 192 - Errors: Null Dereference (Fortify)</td>
<td>6.00</td>
<td></td>
</tr>
</tbody>
</table>

- **No Vulnerabilities**
- **Severity 1**
- **Severity 2**
- **Severity 3**
- **Severity 4**
- **Severity 5**
- **Severity 6**
- **Severity 7**
- **Severity 8**
- **Severity 9**
- **Severity 10**

Copyright 2010, Applied Visions, Inc., Secure Decisions Division
StandardWrapperValve.java

Severity Score: 8.25
Vulnerabilities: 2
Minimum Severity: 6
Maximum Severity: 10

Vulnerabilities | Severity Score | Distribution | Objects
---|---|---|---
2 | 8.25 | | StandardWrapperValve
2 | 8.25 | |
1 | 10.00 | |
1 | 6.00 | |

Bug Created

Created bug #15 and assigned it to remm

OK
Starting point
33,895 vulnerabilities

2 clicks later ...
227 vulnerabilities
Benefits

• Increased vulnerability coverage through the integration of multiple tools
• Overview of large number of vulnerabilities
• Visual prioritization of vulnerabilities
• Traceability of developer responsibility
• Remediation via integration with SDLC
Benefits

- Increased vulnerability coverage through the integration of multiple tools
- Overview of large number of vulnerabilities
- Visual prioritization of vulnerabilities
- Traceability of developer responsibility
- Remediation via integration with SDLC

Enhances the coverage & speed for detection & remediation of vulnerabilities
Visual Analytics

Analytics ← → Visualization
Future Research

- Visualization + Data Analytics
- Event Graphs: Time + Relationships
- Collaborative visualization
- Real-time event stream monitoring

- Situational understanding
- Malware analysis
- Digital forensic analysis
Visualization Toolkits

Prefuse
http://prefuse.org/

Processing
http://processing.org/

Flare
http://flare.prefuse.org/

Protovis
http://vis.stanford.edu/protovis/

Google
http://code.google.com/apis/visualization/