Assessing Vulnerabilities in Software Systems: A Quantitative Approach

Yashwant K Malaiya
BMAC Sept 17, 07

Presentation Outline

- Motivation and Background
- Vulnerability data and AML Vulnerability Discovery Model
  - Fitting HTTP Web servers’ and Browsers’ datasets
  - Measuring Prediction Capabilities
- Fitting categorized vulnerabilities
  - Categories
  - Severity Levels
  - Analyzing categorized vulnerabilities
- Conclusions and Future Work
Motivation

- Security maintenance management
  - Security patch scheduling
  - Estimating the number of vulnerabilities
- Growing risk caused by software vulnerabilities
  - Loss estimation
  - Ownership cost estimation

Research Objectives

- This research look into ways to estimate the number of potential vulnerabilities
- Providing users with:
  - Improving risk assessment
  - Comparison tools
  - Lay a foundation for stable metrics
- Providing developers with
  - Guidelines and tips for better development
  - Testing optimization tools
  - Improving software maintenance resource allocation
Definition and Background

Definition: vulnerability is “a defect which enables an attacker to bypass security measures” [Schultz et al. 1990]

Related work:
- " [Rescorla 2005] and [Ozment 2005] applied some SRGMs to vulnerabilities
- Anderson suggested a model based on thermodynamics analogy
- Arbach et al. proposed a model for incidents

Data Description

- Vulnerability Databases
- Defects Data
  - Vendors
  - Bugzilla, http://www.bugzilla.org
  - Independent Articles
- Market share
  - Google (until mid-2004)
  - Internet counters
- Software sizes
  - Vendors
  - Experts
Vulnerability Datasets

- Operating Systems
  - Windows and Red Hat Linux
- Web-servers
  - Microsoft IIS and Apache
- Web-browsers
  - Internet Explorer and Fire Fox

Vulnerability Discovery Models

- Security and reliability analogy
- VDMs and SRGMs analogy
- Characterizing security
  - Number of potential vulnerabilities
  - Vulnerability discovery rate
The Logistic Model

- A time-based model

\[ \Omega(t) = \frac{B}{B + Ce^{-At}} \]

**The Linear Model**

- Approximation of the logistic model
- The learning phase may be very brief
- Significant number of vulnerabilities are still present and continue to be found
- The dataset is a superimposition of two or more consecutive S-shaped models

\[ \Omega(t) = (S \times t) + k \]
Validating The Logistic Model

Windows 95

Cumulative Vulnerabilities

0 10 20 30 40 50 60 70

Cumulative Vulnerabilities

Actual
AML

Significant

0.9665
69.83216
34.964
52
0.11979
114.25045
0.000528
Solaris 9.0

Cumulative Vulnerabilities

Significant

0.7877
101.8795
69.7045
80
0.26380
99.815
0.000961
Solaris 8.0

Cumulative Vulnerabilities

Significant

0.1759
118.7516
107.701
95
0.11076
126.32
0.000526
Solaris 7.0

Cumulative Vulnerabilities

Significant

0.425800
44.98534
31.81382
31
0.53497
139.045
0.002014
Red Hat

Fedora

Cumulative Vulnerabilities

Significant

0.99456
84.82064
39.6227
65
0.29531
166.735
0.169437
R H Linux

7.1

Cumulative Vulnerabilities

Significant

0.999956
96.21667
36.19688
75
0.13973
121.235
0.000855
R H Linux

6.2

Cumulative Vulnerabilities

Insignificant

0
168.613
267.173
140
0.44439
168.613
Win NT 4.0

Cumulative Vulnerabilities

Significant

0
45.748
159.8135
1
Win 2000

Cumulative Vulnerabilities

Significant

0
166.735
45.748
132
0.32092
166.735
Win 98

Cumulative Vulnerabilities

Insignificant

0
45.748
168.613
132
0.32092
168.613
Win 95

Fit Result

Significant

Insignificant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Fit Result

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Fit Result

Insignificant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Fit Result

Insignificant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Significant

Goodness of fit test-AML

<table>
<thead>
<tr>
<th>System</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>DF</th>
<th>$\chi^2$</th>
<th>$\chi^2_{critical}$ (5%)</th>
<th>P-value</th>
<th>Fit Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Win 95</td>
<td>0.001652</td>
<td>49.5512</td>
<td>1.32092</td>
<td>132</td>
<td>45.748</td>
<td>159.8135</td>
<td>1</td>
<td>Significant</td>
</tr>
<tr>
<td>Win 98</td>
<td>0.000522</td>
<td>92.6789</td>
<td>0.10233</td>
<td>92</td>
<td>107.001</td>
<td>115.3898</td>
<td>0.135785</td>
<td>Significant</td>
</tr>
<tr>
<td>Win XP</td>
<td>0.000232</td>
<td>280.513</td>
<td>0.09994</td>
<td>59</td>
<td>54.6261</td>
<td>77.93052</td>
<td>0.637299</td>
<td>Significant</td>
</tr>
<tr>
<td>Win NT 4.0</td>
<td>0.000317</td>
<td>184.78</td>
<td>0.44439</td>
<td>140</td>
<td>267.173</td>
<td>168.613</td>
<td>0</td>
<td>Insignificant</td>
</tr>
<tr>
<td>Win 2000</td>
<td>0.0001096</td>
<td>391.984</td>
<td>0.0386</td>
<td>69</td>
<td>95.1500</td>
<td>100.7486</td>
<td>0.104079</td>
<td>Significant</td>
</tr>
<tr>
<td>R H Linux 6.2</td>
<td>0.000855</td>
<td>121.235</td>
<td>0.13973</td>
<td>75</td>
<td>36.1988</td>
<td>96.21667</td>
<td>0.999956</td>
<td>Significant</td>
</tr>
<tr>
<td>R H Linux 7.1</td>
<td>0.169437</td>
<td>166.735</td>
<td>0.29531</td>
<td>65</td>
<td>39.6227</td>
<td>84.82064</td>
<td>0.99456</td>
<td>Significant</td>
</tr>
<tr>
<td>Red Hat Fedora</td>
<td>0.002014</td>
<td>139.045</td>
<td>0.53497</td>
<td>31</td>
<td>31.81382</td>
<td>44.98534</td>
<td>0.425800</td>
<td>Significant</td>
</tr>
<tr>
<td>Solaris 7.0</td>
<td>0.000526</td>
<td>126.32</td>
<td>0.11076</td>
<td>95</td>
<td>107.701</td>
<td>118.7516</td>
<td>0.1759</td>
<td>Significant</td>
</tr>
<tr>
<td>Solaris 6.0</td>
<td>0.000961</td>
<td>99.815</td>
<td>0.26380</td>
<td>80</td>
<td>69.7045</td>
<td>101.8795</td>
<td>0.7877</td>
<td>Significant</td>
</tr>
<tr>
<td>Solaris 9.0</td>
<td>0.000528</td>
<td>114.25045</td>
<td>0.11979</td>
<td>52</td>
<td>34.964</td>
<td>69.83216</td>
<td>0.9665</td>
<td>Significant</td>
</tr>
</tbody>
</table>
The Effort-Based Model

- Data needed: The total effort spent using the software system
  - Percentage of market share
  - Global internet population
  - Effort is calculated: \[ E = \sum_{i=0}^{n} (U_i \times P_i) \]

- The effort is mapped to the vulnerability datasets by eliminating the time

- The effort/vulnerabilities data is fitted to the model:
  \[ \Omega(E) = B(1 - e^{-\lambda_\Omega E}) \]
Web Servers Datasets Overview

- Apache and IIS market share (90.6%)

<table>
<thead>
<tr>
<th></th>
<th>Apache</th>
<th>IIS</th>
<th>SJSWS (SunOne)</th>
<th>Zeus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Share</td>
<td>69.7%</td>
<td>20.92%</td>
<td>2.53%</td>
<td>0.78%</td>
</tr>
<tr>
<td>Version</td>
<td>1.x</td>
<td>2.x</td>
<td>4.0</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>Up to 6.1</td>
<td>Up to 4.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vulnerabilities</td>
<td>58</td>
<td>45</td>
<td>85</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Web servers Datasets Goodness of Fit Test

- Software Sys | Fit
- Apache 1.x   | Significant
- Apache 2.x   | Significant
- IIS 4.x      | Significant
- IIS 5.x      | Significant
### Web Browsers Datasets

#### Overview

<table>
<thead>
<tr>
<th>Web Browser</th>
<th>IE Market Share</th>
<th>Firefox Market Share</th>
<th>Mozilla Market Share</th>
<th>Safari Market Share</th>
<th>Other Market Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Share</td>
<td>85.2%</td>
<td>9.6%</td>
<td>0.37%</td>
<td>3.06%</td>
<td>1.77%</td>
</tr>
<tr>
<td>Vulnerabilities</td>
<td>265</td>
<td>94</td>
<td>38</td>
<td>24</td>
<td>N/A</td>
</tr>
<tr>
<td>Release Date</td>
<td>Aug 95</td>
<td>Sep 02</td>
<td>Dec 98</td>
<td>Jan 03</td>
<td>N/A</td>
</tr>
</tbody>
</table>

#### Browser Statistics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>IE</td>
<td>88.9%</td>
<td>86.63%</td>
<td>85.82%</td>
<td>83.05%</td>
</tr>
<tr>
<td>Firefox</td>
<td>4.58%</td>
<td>8.69%</td>
<td>11.23%</td>
<td>12.93%</td>
</tr>
</tbody>
</table>

### Web Browsers Datasets

#### Goodness of Fit Test

<table>
<thead>
<tr>
<th>Software Sys</th>
<th>Significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>IE 4.0</td>
<td>Significant</td>
</tr>
<tr>
<td>IE 5.0</td>
<td>Significant</td>
</tr>
<tr>
<td>IE 6.0</td>
<td>Insignificant</td>
</tr>
<tr>
<td>FFox1.0</td>
<td>Insignificant</td>
</tr>
<tr>
<td>FFox1.5</td>
<td>Significant</td>
</tr>
</tbody>
</table>
Web Browsers Datasets
Goodness of Fit Test

Goodness of Fit Results
(Effort-Based Model)

<table>
<thead>
<tr>
<th></th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apache</td>
<td>Significant</td>
</tr>
<tr>
<td>IIS</td>
<td>Significant</td>
</tr>
</tbody>
</table>
Goodness of Fit Results

- AML model fits the datasets
  - The superposition effect: Consecutive S-shaped models because of shared vulnerabilities
- Effort-based model fits the datasets

Prediction Approaches

- Dynamic Approach (Fitting data to VDMs)
- Static Approach (Using Vulnerability Density of a similar system)

Recommended Combined Approach

Software Age
Vulnerabilities Taxonomy

- **Input Validation Error (IVE) (Boundary condition error (BCE), Buffer overflow (BOF))**: Include failure to verify the incorrect input and read/write involving an invalid memory address.
- **Access Validation Error (AVE)**: These vulnerabilities cause failure in enforcing the correct privilege for a user.
- **Exceptional Condition Error Handling (ECHE)**: May arise due to failures in responding to unexpected data or conditions.
- **Environmental Error (EE)**: Triggered by specific conditions of the computational environment.

Vulnerabilities Taxonomy

- **Configuration Error (CE)**: These vulnerabilities result from improper system settings.
- **Race Condition Error (RC)**: These are caused by the improper serialization of the sequences of processes.
- **Design Error (DE)**: These are caused by improper design of the software structure.
- **Others**: Includes vulnerabilities that do not belong to the types listed above, sometimes referred to as nonstandard.
Windows XP

- IVE-BO: 34%
- ECHE: 16%
- DE: 17%
- IVE-BCE: 10%
- IVE-other: 12%
- Other: 6%
- AVE: 6%

Goodness of Fit Results

Windows 2000

<table>
<thead>
<tr>
<th>Vulnerability Type</th>
<th>Parameters</th>
<th>$X^2$</th>
<th>P-value</th>
<th>$X^2_{\text{Critical}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVE</td>
<td>0.002121</td>
<td>23.181</td>
<td>0.279309</td>
<td>1</td>
</tr>
<tr>
<td>ECHB</td>
<td>0.001026</td>
<td>51.954</td>
<td>0.189765</td>
<td>0.99956</td>
</tr>
<tr>
<td>BOF</td>
<td>0.000676</td>
<td>93.862</td>
<td>0.495049</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>0.001791</td>
<td>54.358</td>
<td>0.343025</td>
<td>0.99446</td>
</tr>
</tbody>
</table>

Windows XP

<table>
<thead>
<tr>
<th>Vulnerability Type</th>
<th>Parameters</th>
<th>$X^2$</th>
<th>P-value</th>
<th>$X^2_{\text{Critical}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVE</td>
<td>0.001035</td>
<td>63.595</td>
<td>2.0119</td>
<td>1</td>
</tr>
<tr>
<td>ECHB</td>
<td>0.000700</td>
<td>92.864</td>
<td>0.674</td>
<td>1</td>
</tr>
<tr>
<td>BOF</td>
<td>0.000808</td>
<td>93.578</td>
<td>0.346</td>
<td>1</td>
</tr>
<tr>
<td>DE</td>
<td>0.002854</td>
<td>28.984</td>
<td>0.1958</td>
<td>1</td>
</tr>
</tbody>
</table>
OS: Vulnerability Categories

Severity Levels

- Common Vulnerabilities Scoring System: A 10 points system to measure the risk of the vulnerabilities
  - High (7-10)
  - Medium (4-6.99)
  - Low (1-3.99)
### Modeling Vulnerabilities by Severity

#### Apache

<table>
<thead>
<tr>
<th>Vulnerability Severity Level</th>
<th>Parameters</th>
<th>$X^2$</th>
<th>P-value</th>
<th>$X^2$ Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>.00156</td>
<td>27.00</td>
<td>1.00</td>
<td>42.1</td>
</tr>
<tr>
<td>Low</td>
<td>.00248</td>
<td>18.00</td>
<td>1.76</td>
<td>15.7</td>
</tr>
</tbody>
</table>

#### IIS

<table>
<thead>
<tr>
<th>Vulnerability Severity Level</th>
<th>Parameters</th>
<th>$X^2$</th>
<th>P-value</th>
<th>$X^2$ Critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>.00176</td>
<td>38</td>
<td>.999</td>
<td>28.2</td>
</tr>
<tr>
<td>Low</td>
<td>.00127</td>
<td>77.9</td>
<td>1.21</td>
<td>53.5</td>
</tr>
</tbody>
</table>

#### Graphical Representation

- **X-axis**: Time (Jan-93 to Jul-02)
- **Y-axis**: Vulnerabilities
- **Legend**: High, Low

---

The graphs show the trend of vulnerabilities over time for Apache and IIS, with distinct lines for high and low severity levels.
**Categories vs. Severity level**

### Apache

<table>
<thead>
<tr>
<th>Category</th>
<th>AVE</th>
<th>DE</th>
<th>ECH</th>
<th>IVE B</th>
<th>IVE BCE</th>
<th>IVE other</th>
<th>RC</th>
<th>EE</th>
<th>CE</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>10</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>35</td>
</tr>
<tr>
<td>Medium</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Low</td>
<td>3</td>
<td>17</td>
<td>12</td>
<td>3</td>
<td>1</td>
<td>14</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>4</td>
<td>68</td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>23</td>
<td>16</td>
<td>13</td>
<td>2</td>
<td>24</td>
<td>2</td>
<td>4</td>
<td>14</td>
<td>5</td>
<td>109</td>
</tr>
</tbody>
</table>

### Windows 2000

<table>
<thead>
<tr>
<th>Category</th>
<th>AVE</th>
<th>DE</th>
<th>ECH</th>
<th>IVE B</th>
<th>IVE BCE</th>
<th>IVE other</th>
<th>RC</th>
<th>EE</th>
<th>CE</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>10</td>
<td>29</td>
<td>12</td>
<td>52</td>
<td>18</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>135</td>
</tr>
<tr>
<td>Medium</td>
<td>6</td>
<td>15</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>Low</td>
<td>6</td>
<td>26</td>
<td>32</td>
<td>9</td>
<td>3</td>
<td>16</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>70</td>
<td>46</td>
<td>66</td>
<td>21</td>
<td>25</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>3</td>
<td>268</td>
</tr>
</tbody>
</table>

**Results Summary**

- **Buffer overflow**
  - Linked to high severity vulnerabilities
- **Exceptional Control Handling Error**
  - Linked to low severity vulnerabilities
- **Input Validation errors**
  - Vulnerabilities of this type other than Buffer overflow and boundary condition error has also shown to be linked to low severity vulnerabilities
- **Design Error vulnerabilities**
  - Found to be a significant category of which a significant proportion is usually of high severity
Summary

- The AML was validated on systems, web servers and web browsers
- The Effort-based model was validated on web server’s datasets
- The prediction capabilities were tested
- Several estimation approaches were suggested

Summary

- Constrained version of AML model was suggested
- Prediction capability is comparable to SRGMs
- Categorized vulnerabilities has been analyzed
- The AML was fitted to categorized data
- Some severity levels has shown strong relationship with some vulnerability categories
Future Work

- White Box analysis
  - Identifying vulnerable modules
  - Analyzing the impact of the code reuse
- Studying the impact of patches
  - what happens after a major patch is applied (e.g. SP1, SP2)?

Future Work

- Evaluating the vulnerabilities economical impact
  - Incentives for vulnerabilities finders
- Real-time quantitative risk assessment using
  - Vulnerabilities and incidents
- Using Testing-oriented categorization of vulnerabilities
Beta test version effect

Multiple Vulnerability Discovery Model (MVDM)

- Hypothesis of MVDM
Multiple Vulnerability Discovery Model (MVDM) – Cont.

- Cumulative MVDM

\[ \Omega(t) = \frac{B}{B C e^{-A t}} + 1 + \alpha B' C' e^{-A' B'(t-t^*)} + 1 \]

One-hump MVDM

<One-Hump MVDM Rate>  <One-Hump MVDM>
### Seasonality of Vulnerabilities

<table>
<thead>
<tr>
<th>Month</th>
<th>Seasonal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>March</td>
<td>1.213</td>
</tr>
<tr>
<td>April</td>
<td>1.846</td>
</tr>
<tr>
<td>May</td>
<td>1.123</td>
</tr>
<tr>
<td>June</td>
<td>1.888</td>
</tr>
<tr>
<td>July</td>
<td>0.008</td>
</tr>
<tr>
<td>August</td>
<td>-0.017</td>
</tr>
<tr>
<td>September</td>
<td>2.226</td>
</tr>
<tr>
<td>October</td>
<td>1.141</td>
</tr>
<tr>
<td>November</td>
<td>-0.200</td>
</tr>
<tr>
<td>December</td>
<td>2.068</td>
</tr>
<tr>
<td>January</td>
<td>-1.059</td>
</tr>
<tr>
<td>February</td>
<td>-2.122</td>
</tr>
</tbody>
</table>

- ARIMA (AutoRegressive Integrated Moving Average)

### Recent Conference Papers