POSTER: Understanding the Hidden Cost of Software Vulnerabilities: Measurements and Predictions

Afsah Anwar
University of Central Florida
 afsahanwar@knights.ucf.edu

Aminollah Khormali
University of Central Florida
 aminkhormali@knights.ucf.edu

Aziz Mohaisen
University of Central Florida
mohaisen@ucf.edu

ABSTRACT
In this work, we study the hidden cost of software vulnerabilities reported in the National Vulnerability Database (NVD) through stock price analysis. We perform a high-fidelity data augmentation to ensure data reliability for estimating vulnerability disclosure dates as a baseline for assessing software vulnerabilities’ implication. We further build a model for stock price prediction using the NARX Neural Network model to estimate the effect of vulnerability disclosure on the stock price. Compared to prior work, which relies on linear regression models, our approach is shown to provide better accuracy. Our analysis shows that the effect of vulnerabilities on vendors varies, and greatly depends on the specific industry.

KEYWORDS
Vulnerability Economics; Prediction; NVD

ACM Reference Format:

1 INTRODUCTION
Vulnerabilities are defects in software products, exposing them and users to risk alike. To deal with vulnerabilities, vendors incur costs in the form of developer-hours spent fixing them and re-deploying the fixes; a study by the National Institute of Standards and Technology (NIST) estimated that the US economy about $60 Billion USD every year for patches development and redistribution, systems re-deployment, and direct productivity loss [16]. In 2012, Knight Capital, a fintech company, lost $400 Million USD because of a bug in their code; the company bought shares at the ask price and sold them at the bid price [17]. Losses from WannaCry (2017), a ransomware attack in over 150 countries affecting more than 100,000 groups, is estimated at $4 Billion USD [4]. Virus attacks, such as Love Bug (2000), SirCam (2001), Nimda (2001), and CodeRed (2001), have had an impact of $8.75 Billion, $1.25 Billion, $1.5 Billion and $2.75 Billion USD, respectively [1]. The direct losses, however, are not the only cost of vulnerabilities, and companies incur additional hidden costs that we attempt to assess in this work.

Figure 1: Dataset Creation Flow. Desc. stands for vulnerability description, Ref Link is the link referring to details of the vulnerability, Pub. Date is the Published Date, CVSS is Common Vulnerability Scoring System metrics, CWE is the Common Weakness Enumeration identifier, PDD is the Public Disclosure Date, approximated as the minimum of the dates gathered from the links corresponding to a vulnerability, and VHSP is the Vendor Historical Stock Price.

2 METHODOLOGY
Using the information available on the NVD, the goal of this study is to track the public disclosure date of vulnerabilities and capture their impact on vendors stock market valuation.

2.1 Data and Data Augmentation
Figure 1 summarizes, at a high-level, the flow of data creation, from the source of data to the final dataset.

National Vulnerability Database (NVD). is a vulnerability database maintained by the National Institute of Standards and Technology (NIST) and contains the vulnerabilities reported to MITRE [2].

The NVD includes the following elements for each vulnerability: the CVE-ID, vendor, product, Common Vulnerability Scoring System (CVSS) label, published date, Common Weakness Enumeration Identifier (CWE-ID) [3], description, and reference links.

Data Preprocessing and Augmentation. We use the vulnerabilities reported in the year 2016, and limit our analysis to the severe ones. To assess the impact of intrinsic vulnerabilities in software, rather than inherited vulnerabilities due to third-party libraries, we discard vulnerabilities with “library” in their description. We use the links from the NVD to scrape through the web and label public disclosure date for each of the vulnerabilities by calculating the minimum of the dates (where multiple dates are available). For 1,262 links that prevent automated scraping, we manually visited...
2.2 Assessing Vulnerability’s Impact

To assess the effect of vulnerability on the day for which the vulnerability was published, we look for the stock value on that particular date. For all dates with disclosed vulnerabilities for which the stock data is available and the value on the last operating day, we look for the stock value on that day. The value of the vendor’s stock achieved on the given day. The adjusted close attribute corresponds to the dividends and splits since that day. The volume is the number of shares traded on the given day.

Press. As a baseline for comparison with our results based on the approach used in the literature, we sample vulnerabilities reported in the press. We search for “software vulnerabilities in 2017” in Forbes and ZDNet, and list four vulnerabilities for comparison.

3 PREDICTION

Normalization. Each feature, open, close, high, low, volume, and adjacent close, is transformed into feature vector \( z = \frac{x - \bar{x}}{\sigma} \), where \( \bar{x} \) and \( \sigma \) are the mean and standard deviation of the original feature vector \( x \), respectively. The resulting features are then fed into the nonlinear auto-regressive neural network with exogenous factors (NARX) to predict the stock value of vendors.
Table 1: Results for each Vendor. Vul. stands for vulnerability count and \( OAR_1 \), \( OAR_2 \), and \( OAR_3 \) stand for the average effect at day 1, 2, and 3 (percent), respectively. ▲ indicates that the vulnerabilities had no overall impact on vendor’s stock value while ▼ indicates that the stock of the vendor were impacted, overall.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>MSE</th>
<th>Vul.</th>
<th>( OAR_1 )</th>
<th>( OAR_2 )</th>
<th>( OAR_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adobe</td>
<td>5.9E-4</td>
<td>494</td>
<td>▼0.65</td>
<td>▼0.37</td>
<td>▼0.50</td>
</tr>
<tr>
<td>Advantech</td>
<td>9.5E-4</td>
<td>9</td>
<td>▲0.61</td>
<td>▲0.89</td>
<td>▲0.96</td>
</tr>
<tr>
<td>Apache</td>
<td>9.9E-4</td>
<td>37</td>
<td>▲0.60</td>
<td>▲0.98</td>
<td>▲1.17</td>
</tr>
<tr>
<td>Apple</td>
<td>2.8E-4</td>
<td>154</td>
<td>▲0.41</td>
<td>▲0.75</td>
<td>▲1.03</td>
</tr>
<tr>
<td>Atlassian</td>
<td>9.7E-3</td>
<td>4</td>
<td>▼3.85</td>
<td>▼3.86</td>
<td>▼3.12</td>
</tr>
<tr>
<td>Cisco</td>
<td>2.3E-3</td>
<td>111</td>
<td>▲0.10</td>
<td>▲0.33</td>
<td>▲0.42</td>
</tr>
<tr>
<td>Citrix</td>
<td>2.4E-3</td>
<td>9</td>
<td>▲0.14</td>
<td>▲0.01</td>
<td>▲0.57</td>
</tr>
<tr>
<td>Facebook</td>
<td>1.1E-3</td>
<td>6</td>
<td>▲0.13</td>
<td>▲0.33</td>
<td>▲0.45</td>
</tr>
<tr>
<td>Fortinet</td>
<td>4.5E-3</td>
<td>7</td>
<td>▲0.37</td>
<td>▲0.19</td>
<td>▲0.92</td>
</tr>
<tr>
<td>GC</td>
<td>5.8E-4</td>
<td>3</td>
<td>▲0.12</td>
<td>▼0.58</td>
<td>▼0.39</td>
</tr>
<tr>
<td>Google</td>
<td>7.6E-4</td>
<td>410</td>
<td>▼0.08</td>
<td>▼0.21</td>
<td>▼0.08</td>
</tr>
</tbody>
</table>

CRITICAL. Vulnerabilities involving unauthorized accesses have a higher cost, seen in their effect on the stock price. (3) Denial of Service attacks unrelated to confidentiality do not impact stock.

For the vulnerabilities gathered from the press, we followed the same steps and observe that these vulnerabilities have an adverse effect on vendor stock price in almost every case.

Discussion. Prior works show the impact of vulnerabilities using Cumulative Abnormal Rate (CAR), which aggregates AR’s on different days. However, CAR does not effectively capture the impact of a vulnerability. For example, CAR would indicate no effect if the magnitude (upward) of one or more days analyzed negate the magnitude (downward) of other days. Also, our results through a rigorous analysis are statistically significant. One main shortcoming of the prior work, however, is that it overlooks analyzing the cost based on sectors of the software industry. In this study we found a significant effect of vulnerabilities on a given day and limited ourselves to the third day after the release of the vulnerability in order to minimize the impact of other factors. Eliminating effects of other factors and measuring long-term losses is an open question.

5 RELATED WORK

While Hovav and D’Archy [12] have shown that market shows no signs of significant negative reaction due to vulnerabilities, Telang et al. [18] show that a vendor on average loses 0.6% of its stock value due to vulnerabilities. Goel et al. [10] also pointed out that security breaches have an adverse impact on the market value of a vendor. Bose et al. [5] show that each phishing alert leads to a loss of market capitalization that is at least US$ 411 million for a firm.


Table 2: Per industry stock impact likelihood analysis.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Likeliness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software</td>
<td>Highly Likely</td>
</tr>
<tr>
<td>Consumer Products</td>
<td>Highly Likely</td>
</tr>
<tr>
<td>Finance</td>
<td>Highly Likely</td>
</tr>
<tr>
<td>Security</td>
<td>Equally Likely</td>
</tr>
<tr>
<td>Electronics &amp; Hardware</td>
<td>Equally Likely</td>
</tr>
<tr>
<td>Conglomerate</td>
<td>Less Likely</td>
</tr>
<tr>
<td>Device</td>
<td>Less Likely</td>
</tr>
<tr>
<td>Networking</td>
<td>Less Likely</td>
</tr>
</tbody>
</table>

6 CONCLUSION AND FUTURE WORK

We perform an empirical analysis on vulnerabilities from NVD and look at their effect on vendor’s stock price. Our results show that the effect is industry-specific, and depends on the severity of the reported vulnerabilities. We also compare the results with the vulnerabilities found in popular press: while both vulnerabilities affect the vendor’s stock, vulnerabilities reported in the media have a much more adverse effect. En route, we also design a model to predict the stock price with high accuracy.

Acknowledgement. This work is supported in part by NSF grant CNS-1809000 and NRF grant NRF-2016K1A1A2912757.

REFERENCES