An Empirical Study of Security Problem Reports in Linux Distributions

Prasanth Anbalagan
panbala@ncsu.edu
Department of Computer Science
North Carolina State University
Raleigh, USA 27695

Mladen Vouk
vouk@ncsu.edu
Department of Computer Science
North Carolina State University
Raleigh, USA 27695

ABSTRACT
Existing studies on problem reports in open source projects focus primarily on the analysis of the general category of problem reports, or limit their attention to observations on the number of security problem reports. To evaluate the security of a project, it is necessary to know not only how many security problem reports are logged but also how many are exploited in the field, which problem reports and how quickly they are corrected, etc. In this paper, we study security of a project, it is necessary to know how often one exposure to and correction of such reports. To evaluate the number of security problem reports or limit their attention to observations on number of security problems reported. We compute quantitative measures to analyse which of those security problem reports and how promptly they are corrected and discuss the security of each of the projects.

1. INTRODUCTION
Existing work like [2, 5] focuses primarily on the analysis of the general category of problem reports or limit their attention to observations on number of security problems reported in open source projects. Although, security problem reports and failures, are a subset of the general category of problem reports and failures [6], our earlier work [3] showed that security problem reports have characteristics different from the general category of problem reports in terms of exposure to and correction of such reports. To evaluate the security of a project, it is necessary to know how often one is exposed to security problem reports, how many are exploited in the field, which of those problem reports and how quickly they are corrected, etc. In this paper, we study security problem reports from eight releases of Fedora, nine releases of Ubuntu, four releases of RedHat Enterprise Linux (RHEL) and two releases of Suse Linux distributions, analyse and discuss which type of problem reports and how frequently they are reported, and how promptly they are corrected. Overall, Fedora and Suse show good results with high and medium severity security problem reports resolved without a backlog. On the other hand, RHEL and Ubuntu show less positive results with presence of backlogs.

2. DATA
Table 1 shows the number of problem reports and installations running more recent releases of Fedora. Rows High, Medium, Low show the number of security problems under each category and their percentage in the total number of problems PR. Tables 2, 3 and 4 show similar statistics for Ubuntu, RHEL and Suse. In this paper, we consider system usage to analyse security from the perspective of the system as a whole, i.e., usage of the total number of systems in operation during a given time, the total operation time as well as downtime for each linux installation, etc. we calculate the inservice time i.e., product of the total number of systems (installations) and the time they have been operational as in [7]. The inservice time approximates Fedora usage as a whole, e.g., the usage of all the ni systems during a particular week. The Fedora download statistics is only available from Fedora release 6 onwards. Therefore our analyses on Fedora as a whole have been done on Fedora releases 6, 7 and 8. As for the linux distributions, we focus our analysis on Ubuntu, Suse and RHEL running on an individual system (usage of one system during a particular week).

3. METRICS (MTTPR AND MTTPC)
Figure 2 shows a sample window of three periods encompassing four problem reports. The problem reports are marked with the time when the problems were reported (Ri) and fixed(Fi). In Figure 2, Xi gives the time to problem report (TTPR), the amount of time system is operational without a problem being reported and Yi gives the time to problem correction (TTPC), the amount of time taken to fix a problem report. Mean Time To Problem Report(MTTPR) is the mean or average of all times to problem report during a particular interval. This is analogous to mean-time-to-failures which is the average amount of time a system is operational without failures during a particular time interval. In the period i-1, the Mean Time to Problem Report would be estimated as the average of X1 and X2. Mean Time To Problem Correction(MTTPC) is defined as the average of all times to problem correction (TTPC) during a particular interval. For period i, the Mean Time to Problem Correction is the

*This work is supported by the U.S. Army Research Office (ARO) under grant W911NF-08-1-0105 managed by NCSU Secure Open Systems Initiative (SOSI).
1http://fedoraproject.org/

2http://www.ubuntu.com/
3http://www.redhat.com/
4http://www.novell.com/linux/
Table 1: Fedora Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>F 1</th>
<th>F 2</th>
<th>F 3</th>
<th>F 4</th>
<th>F 5</th>
<th>F 6</th>
<th>F 7</th>
<th>F 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security (PRs)</td>
<td>31</td>
<td>128</td>
<td>82</td>
<td>154</td>
<td>201</td>
<td>194</td>
<td>92</td>
<td>26</td>
</tr>
<tr>
<td>Non-security (PRns)</td>
<td>3141</td>
<td>4124</td>
<td>7528</td>
<td>7644</td>
<td>7805</td>
<td>7877</td>
<td>5488</td>
<td>4490</td>
</tr>
<tr>
<td>Total (PRs)</td>
<td>3172</td>
<td>4206</td>
<td>7656</td>
<td>7798</td>
<td>8006</td>
<td>8071</td>
<td>5560</td>
<td>4516</td>
</tr>
<tr>
<td>% of (PRns)</td>
<td>0.98</td>
<td>3.01</td>
<td>1.07</td>
<td>1.97</td>
<td>2.51</td>
<td>2.4</td>
<td>1.65</td>
<td>0.58</td>
</tr>
</tbody>
</table>

- **High**: 13 (42%), 54 (42%), 46 (56%), 63 (41%), 74 (37%), 51 (26%), 23 (25%), 8 (31%)
- **Medium**: 15 (48%), 62 (48%), 26 (32%), 73 (47%), 93 (46%), 113 (58%), 60 (65%), 16 (62%)
- **Low**: 3 (10%), 12 (9%), 10 (12%), 18 (12%), 34 (17%), 30 (16%), 9 (10%), 2 (8%)

Table 2: Ubuntu Statistics

<table>
<thead>
<tr>
<th>Statistics</th>
<th>4.10</th>
<th>5.04</th>
<th>5.10</th>
<th>6.04</th>
<th>6.10</th>
<th>7.04</th>
<th>7.10</th>
<th>8.04</th>
<th>8.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Security (PRs)</td>
<td>107</td>
<td>51</td>
<td>54</td>
<td>34</td>
<td>58</td>
<td>186</td>
<td>336</td>
<td>192</td>
<td>87</td>
</tr>
<tr>
<td>Non-security (PRns)</td>
<td>4177</td>
<td>5145</td>
<td>11248</td>
<td>6947</td>
<td>9029</td>
<td>9482</td>
<td>12085</td>
<td>9876</td>
<td>3388</td>
</tr>
<tr>
<td>Total (PRs)</td>
<td>4284</td>
<td>5196</td>
<td>11302</td>
<td>6981</td>
<td>9087</td>
<td>9668</td>
<td>12421</td>
<td>10068</td>
<td>3475</td>
</tr>
<tr>
<td>% of (PRns)</td>
<td>2.50</td>
<td>0.98</td>
<td>0.48</td>
<td>0.49</td>
<td>0.64</td>
<td>1.92</td>
<td>2.71</td>
<td>1.91</td>
<td>2.57</td>
</tr>
</tbody>
</table>

- **High**: 47 (44%), 23 (45%), 24 (44%), 16 (47%), 22 (38%), 52 (28%), 134 (40%), 70 (36%), 19 (22%)
- **Medium**: 38 (36%), 23 (45%), 23 (45%), 16 (47%), 32 (55%), 108 (58%), 178 (53%), 107 (56%), 56 (64%)
- **Low**: 22 (20%), 5 (10%), 7 (13%), 2 (6%), 4 (7%), 26 (14%), 24 (7%), 15 (8%), 12 (14%)

Figure 1: Problem Report/Resolution Time Intervals
total correction time (sum of time periods Y2 and Y3, the times spent to fix problem reports 3 and 4) divided by two.

4. RATES

In our analysis, we estimate problem exposure rates, problem correction rates, and problem exploit rates. We define problem exposure rate as the number of unresolved or open problem reports a system has in the time period of interest. Since an end-user is exposed to not only new problems disclosed during a particular time period, but also unresolved problem reports from the previous time period which could also potentially be exploited by any malicious user, we consider problem reports reported during the current time period as well as unresolved problem reports from the previous time periods. This is different from the exposure to only the unique number of problem reports reported during a particular time period. We define problem correction rate as the number of security problems resolved in the time period of interest. We define problem exploit rate as the number of security problem reports actually exploited in the field in the time period of interest.

4.1 Problem exposure rates

Let the problem exposure rate \( \lambda_d \) for publicly disclosed security problem reports (inclusive of high, medium and low severity categories) in the time interval \( i \) be estimated as,

\[
\lambda_d = \frac{(n_o)_i + (n_f)_i}{(n_i) \times \Delta t_i}
\]  

(1)

where \((n_o)_i\) represents the number of open problem reports from the previous week, \((n_f)_i\) represents number of problems reported in time interval \( \Delta t_i \) (e.g., given in minutes). Typically, the number of users \((n_i)\) in time interval \( \Delta t_i \) is very large compared to the total number of problem reports considered \(((n_o)_i + (n_f)_i)\) [Table 1]. It is worth noting again that \(\lambda_d\) is different from the inverse of the mean time to problem resolution (MTTPR). The latter reflects arrival time of unique problem reports in the calendar time frame, while the former reflects the number of open reports with respect to the inservice time during time period \( i \). We may wish to use \(\lambda_d\) when discussing the security of the system as a whole, and MTTPR in the context of calendar-based problem correction rate. The problem exposure rates [3] for high \((\lambda_h)\), medium \((\lambda_m)\), and low severity \((\lambda_l)\) security problem reports can be estimated in a similar way as in equation (1). Here, \((n_o)_i\) and \((n_f)_i\) would represent only the problems from each of the respective categories. Table 5 shows the rates for various Linux distributions. Under our assumptions, these rates exhibit behavior and consistency typical of low failure rate systems for which we have the actual total failure data e.g. failure and correction rates of high-quality telecommunication systems [4].

4.2 Problem correction rates

The problem correction rate for publicly disclosed problems can be estimated through the inverse of the Mean Time To Problem Correction (MTTPC), where this time refers only to disclosed security problem reports in time frame, i.e.,

\[
\mu_d = \frac{1}{MTTPC_i}
\]  

(2)
The problem correction rates \( \mu_h \) for high, medium \( \mu_m \), and low severity \( \mu_l \) security problem reports can be estimated in a similar way as in equation (2). Here, the MTITPC is calculated only for problems from each of the respective categories. Problem correction rate appears to be several orders of magnitude larger than the exposure rate.

### 4.3 Problem exploit rates

In practice, data such as the actual field exploits, operational profile of the system, etc., are required to estimate system exploit rate accurately. Information about the actual exploits was not available and the number of problem reports does not tell how many exploits are experienced by an end-user. However empirical results from [1] show on the average the proportion of publicly disclosed and undisclosed problems exploited in the field. We use this to estimate the exploit rates by [7]. Similar to problem exposure rates, we estimate exploit rates for various categories of problem reports. Consider again equation (1). If we assume that \( P_e \) is the probability that exposure to an open problem report results in a field exploit, then

\[
\lambda_d = \frac{1}{n_i} \times \frac{(n_f)_i + (n_u)_i}{\Delta t} \times P_e \quad (3)
\]

\[
\lambda_{ic} = \lambda_d \times P_e \quad (4)
\]

Where \( \lambda_d \) is an estimate of Fedora exploit in the field for publicly disclosed problems (inclusive of high, medium and low severity problem reports). The problem exploit rate for high \( \lambda_{h,c} \), medium \( \lambda_{m,c} \) and low \( \lambda_{l,c} \) severity problem reports can be estimated in a similar way as in (4) using problem exposure rates from the respective categories. In Table 5, we use \( P_e = 0.22 \) for the proportion publicly disclosed problems exploited from [1].

### 5. DISCUSSION

We illustrate the rates computed for Fedora 6, Ubuntu 8.04, RHEL 5, and Suse Linux 10.2 and discuss the results. Similar characteristics and behavior were observed with the other releases of each Linux distribution.

#### 5.1 Fedora Security

We consider problem exposure rates from two perspectives. That of Fedora as a whole with \( n_i \) systems running during a particular week, and that of a particular system that may be under attack. In the first case, there are \( n_i \) systems running Fedora during a particular week. Over these systems, a random Fedora user is exposed to \( (n_f)_i \) unique problems. This includes the \( (n_f)_i \) unique problems reported weekly and \( (n_u)_i \) unique problems unresolved from previous week. Since \( n_i \) is very large (Table 1), the exposure rate of a random Fedora user to unique problems during any particular week week is usually very small. Of course, these problems may reside in all running systems, and in fact be identified more than once among the \( n_i \) systems. Nevertheless, from the perspective of Fedora as a whole, exposure rate to unique problems is relatively low, and a relatively small number of patches should be able to take care of those problems. The exposure rate for a random anonymous system is of the order of 10E-09 security problems per minute. From the perspective of an individual system identified and targeted, about 4 unique problems are reported on an average per week (resulting \( \lambda_d \) and weekly count is shown in Table 5). But from the perspective of Fedora as a whole, these 4 unique problems are reported only from running a very large number of Fedora systems (about 1 to 2 million Fedora systems). The Fedora security problem correction rate shown in Table 5 is thus adequate for the product as a whole. Further, not all of these problems are actually exploited in the field. For example, [1] show that about only 22% of disclosed problems result in exploits.

The exposure rate of a system identified and targeted is about 10E-04 (4 problems per week, Table 5). The Fedora security problem correction rate shown in Table 5 is smaller than the exposure rate of a system that may be under attack. For example, let us consider the problem exposure \( \lambda_d \) and correction rates \( \mu_d \) for disclosed security problems. On an average during a particular week, out of 4 disclosed problems, only half of them are corrected. This leaves the system exposed to approximately 2 additional problems per week. This implies a backlog in the overall problem resolution. Table 5 also shows the correction rate for high, medium, and low severity Fedora security problem reports. We can observe that the correction rate for high and medium severity problems is enough to keep up with the disclosed, but that of low severity problem reports is not. Resolution of such high and medium severity problems with no backlogs is a very good news for Fedora project. If one assumes that only about 22% of the disclosed problems result in exploits [1], the correction rate for high and medium severity problems is good enough to keep up with the exploitable (Table 5). Thus resolution rate of high and medium severity problems shows a positive sign for the project behavior in terms of problem resolution.

Certainly, as an individual user who may not be typically singled out for an attack, one might feel comfortable knowing that there are nearly about 2 million systems operating per week and that chances of a random attack are relatively low. On the other hand, those that may be running high visibility sites, have a higher probability of being targeted. Even in such cases of being singled out, one may feel safe considering
that problem correction rates for high and medium severity problems keep up with problem exposure rates.

5.2 Ubuntu Security
On an average during a particular week, out of 7 disclosed problems, only less than one problem is corrected. This leaves the system exposed to approximately 6 problems per week. This implies a significant backlog in the problem resolution. Table 5 shows the correction rate for high, medium, and low severity Ubuntu security problems. We can observe that the correction rate for low severity problems is enough to keep up with the disclosed, but that of high and medium severity problems are not. Even under the assumption that only about 22% of the disclosed problems result in exploits [1], the resolution rates do not keep up with the exploitable. On the average 33% (resulting from exploitation) of the exploitable high severity problems and 40% of the exploitable medium severity problems reported per week remain unresolved. Since the impact due to these categories is significant, it is necessary to resolve high and medium severity problems much quicker and avoid backlogs.

5.3 Suse Security
The overall resolution rate for the disclosed security problems does not show presence of any backlog. The project behavior for the categories of high, medium and low severity problems show a high resolution rate keeping up with the disclosed problems. But like the other distributions, this category of problems and a backlog due to them may not be significant.

5.4 RedHat Enterprise Linux Security
The overall resolution rate for the disclosed security problems as well as that of high and medium severity problems show presence of backlogs. On the average 20% (resulting from λd and μd in Table 5) of the high severity problems reported per week remain unresolved. But if one considers that only about 22% of the disclosed problems result in exploits [1], then the problem correction rates are good enough to keep up with the exploitable.

<table>
<thead>
<tr>
<th>Rates</th>
<th>Fedora (whole)</th>
<th>Fedora (Individual)</th>
<th>Ubuntu (Individual)</th>
<th>Suse (Individual)</th>
<th>RedHat (Individual)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem exposure rates in security problem reports/min (per week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λd</td>
<td>1.6E-09</td>
<td>3.6E-04 (4)</td>
<td>7.1E-04 (7)</td>
<td>8.2E-05 (1)</td>
<td>2.6E-04 (3)</td>
</tr>
<tr>
<td>λh</td>
<td>2.7E-10</td>
<td>1.1E-04 (1)</td>
<td>2.6E-04 (3)</td>
<td>3.7E-05 (0.4)</td>
<td>5.5E-05 (0.5)</td>
</tr>
<tr>
<td>λm</td>
<td>9.7E-10</td>
<td>1.8E-04 (2)</td>
<td>3.9E-04 (4)</td>
<td>3.7E-05 (0.4)</td>
<td>1.6E-04 (1.0)</td>
</tr>
<tr>
<td>λt</td>
<td>3.4E-10</td>
<td>6.1E-05 (0.6)</td>
<td>5.5E-05 (0.6)</td>
<td>1.2E-05 (0.1) Ad</td>
<td>4.2E-05 (0.4)</td>
</tr>
<tr>
<td>Exploit rates in security problem reports exploited/min (per week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>λde</td>
<td>3.5E-10</td>
<td>7.9E-05 (0.8)</td>
<td>1.6E-04 (2)</td>
<td>1.8E-05 (0.2)</td>
<td>5.7E-05 (0.6)</td>
</tr>
<tr>
<td>λhe</td>
<td>6.0E-11</td>
<td>2.5E-05 (0.3)</td>
<td>5.7E-05 (0.6)</td>
<td>8.2E-06 (0.1)</td>
<td>1.2E-05 (0.1)</td>
</tr>
<tr>
<td>λme</td>
<td>2.1E-10</td>
<td>4.0E-05 (0.4)</td>
<td>8.7E-05 (1)</td>
<td>8.2E-06 (0.1)</td>
<td>3.5E-05 (0.4)</td>
</tr>
<tr>
<td>λte</td>
<td>7.5E-11</td>
<td>1.3E-05 (0.1)</td>
<td>1.2E-05 (0.1)</td>
<td>2.7E-06 (0.03)</td>
<td>9.3E-06 (0.1)</td>
</tr>
<tr>
<td>Correction rates in security problem reports/min (per week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>μd</td>
<td>1.9E-04 (2)</td>
<td>4.3E-05 (0.4)</td>
<td>1.2E-05 (12)</td>
<td>3.8E-05 (0.5)</td>
<td></td>
</tr>
<tr>
<td>μh</td>
<td>3.2E-04 (3)</td>
<td>3.7E-05 (0.4)</td>
<td>8.5E-004 (9)</td>
<td>3.5E-05 (0.4)</td>
<td></td>
</tr>
<tr>
<td>μm</td>
<td>2.0E-03 (20)</td>
<td>5.7E-005 (0.6)</td>
<td>6.5E-04 (7)</td>
<td>5.2E-05 (0.5)</td>
<td></td>
</tr>
<tr>
<td>μt</td>
<td>2.3E-05 (0.2)</td>
<td>5.7E-04 (6)</td>
<td>1.7E-05 (0.2)</td>
<td>4.5E-05 (0.5)</td>
<td></td>
</tr>
</tbody>
</table>

6. CONCLUSIONS
We have analysed security problem reports for various releases of Fedora, Ubuntu, RHEL and SUSE Linux distributions. Using quantitative measures, we have discussed the security of each Linux distribution in terms of exposure to security problem reports during a particular time period, and how promptly they are corrected.

7. REFERENCES