A Hybrid Vulnerability Analysis Tool Using a Risk Evaluation Technique

1Jong-Chul Park, 1Sang-Ann Nam, 1Jae-Pyo Park and 2Yeun-Soo Choo
1Department of IT Policy Management,
2Department of Information Security, Soongsil University, 07027 Seoul, Republic of Korea

Abstract: Recent there have been many efforts to detect and analyze vulnerabilities using diverse analysis tools, removing them at the development stage. However, vulnerability analysis tools are prone to missed detections incorrect detections and over detection which reduces the accuracy of detection. In this study, a vulnerability detection technique is proposed that develops and manages safe applications and can resolve and analyze these problems. Risks due to vulnerabilities are computed and an intelligent vulnerability detection technique is used to improve accuracy and evaluate risks of the final version of the application. This helps the development and execution of safe applications. Through incorporation of tools that use both static and dynamic analysis techniques our proposed technique overcomes weak points at each stage and improves the accuracy of vulnerability detection. Existing vulnerability risk evaluation system only evaluate self-risks while our proposed vulnerability risk evaluation system reflects vulnerability self-risk and detection accuracy in a complex fashion to evaluate relative. Our proposed technique compares and analyzes existing analysis tools such as lists for detections and detection accuracy based on the top 10 items of SANS at CWE. Quantitative evaluation systems for existing vulnerability risks and proposed application vulnerability risks are compared and analyzed. Through, incorporation of tools that use both static analysis and dynamic analysis techniques. We developed prototype analysis tool using our technique to test the application's vulnerability-detection ability and show our proposed technique is superior to existing ones.

Key words: Vulnerability, degree of risk, vulnerability detection, accuracy of detection, application, complex

INTRODUCTION

In order to defend against cyber attacks and minimize damages, many have adopted diverse security solutions for protection yet damage incidents and their scale has still grown. Among the reason for increased cyberattacks is the fact that about 75% are due to software’s weak points (Allodi and Massacci, 2014). The heart bleed bug of open SSL, an open source code which 2/3 of all global web sites are currently using is just one of the typical cases. Microsoft has stated that development cost can be saved for projects that develop security operation systems which consider security issues from the beginning, rather than ones that do not consider them from the beginning and develop security patches as time passes (Anonymous, 2016). Therefore, attempting to find weak points in the development phase of software is a prudent way to reduce cyber attacks and to lower development costs. As techniques to detect vulnerabilities, there are static module based on source codes and dynamic module based on execution systems. However, these methods are prone to false detection, missing detections, over detection and other problems (Mouzarni et al., 2016; Ransbotham and Mitra, 2013). In this study, a vulnerability analysis technique is proposed to help resolve these problems. Our proposed technique is designed with an interaction analysis to increase detection accuracies using existing methods. This increases accuracies of the results as its interaction results re-confirm prior results. Our method has five distinct analysis stages the results that were generated by each stage get analyzed using the dynamic and static modules twice which increases overall accuracy. All results generated in each stage get analyzed first and weights are applied to evaluate the final degree of risks of vulnerability to the application. The existing analysis methods are used to evaluate functions of the proposed vulnerability analysis technique this research compares and analyzes functions using these existing analysis methods. Therefore, our intelligent vulnerability analysis technique is proposed to enhance the accuracy of detection and its application for risk evaluation to evaluate an application vulnerability status by analyzing its weak points. Our proposed technique also has the advantage of automatically executing inspection activities that humans currently perform manually.

Corresponding Author: Jae-Pyo Park, Department of Information Security, Soongsil University, 07027 Seoul, Republic of Korea
MATERIALS AND METHODS

Vulnerability analysis technique: The static analysis technique uses source codes to analyze vulnerability that is inherent in the application without execution. The static analysis technique is not only used to detect vulnerability that is inherent in the completed applications but also used to test patterns, API and others that are confirmed as codes to possibly cause vulnerability from the development phase this allows to prevent occurrence of vulnerability. Therefore, the static analysis technique is used to maintain and minimize repair costs of completed applications. Therefore, it is widely used to maintain and lower repair costs for completed applications (Chen et al., 2013). One of common methods of static analysis techniques is pattern matching i.e., parsing that uses syntax analysis, lexical analysis and others methods. Pattern matching detects vulnerability using string patterns; string patterns use forms of unsafe functions, dangerous variable names, particular code patterns. Since, this method categorizes all patterns that are determined by rules to weak points for their detection, all footnotes are also detected. Therefore, its downside is the high possibility of over detections when compared to dynamic analysis techniques (Fang et al., 2014; Christey, 2011).

Other known static analysis techniques are ‘Parsing’ and ‘Data Flow Analysis. The parsing classifies and interprets source code and source code gets expressed with abstract syntax tree. This parsing tree not only construct program but also analyze meanings. In order to accurately parsing and analyze diverse program, static analysis technique requires one key compiler and compatible parse. The data flow analysis is a traditional compiler technique to resolve ‘Buffer Overflow’ and ‘Format String’ issues. This technique can also be used to detect vulnerability. This technique collects information of possible values that can be calculated from diverse points. Easy way to execute data flow analysis is to set data-flow equations for nodes of control flow graph and calculate from input to output repeatedly until the whole system gets reached to stability (Castro et al., 2006).

The dynamic analysis technique is to analyze whether functions of application run properly or not data gets input into actual running application and outputted real results get analyzed for its vulnerability (Ernst, 2003). It is to test whether errors or unexpected results gets draw or not when the value gets input into application. Simply speaking when ‘1’ and ‘2’ gets input into the Add-function module, this dynamic analysis technique analyzes whether we get ‘3’ as the result or not. At this time, if unintended ‘letter character’ gets input, dynamic analysis technique analyzes whether we get security problems or not. Dynamic analysis techniques can detect functional defects which static analysis technique cannot. Static analysis techniques detect vulnerabilities based on source codes; they cannot detect unexpected results during actual execution. However, dynamic analysis method’s advantage is that all values are input for dummy purging cases this results in detecting all vulnerabilities that developers may not expect and relatively high inspection rates compared to static analysis methods. However, knowledge of input values and the many processes to detect many inputs are downsides in detecting vulnerabilities under this technique (Khurana et al., 2016; Khan and Khan, 2012).

International information for vulnerability management: ISO (International Organization for Standardization) 27005 defines cybersecurity weak points as a group of weaknesses that can be exploited by one or more threats. Also, the internet engineering task force defines weak points as defects and weaknesses from system designs, materializations, operations and management which can be exploited to violate system security policies (ISO., 2016). Weak points are known as security vulnerabilities to actual attacks among many weak points that exist in information assets in systems and solutions. These weak points go through confirmation, categorization, cure and a remedy process which requires routine management (Agawal and Singh, 2013). This type of systematic management of vulnerabilities is known as a vulnerability management system and all different vulnerabilities found from diverse agencies around the world are managed this way.

Application developers and users can make decisions of relevant application’s adjustments and reuse, once it is known how dangerous applications are when they get to know where do managed vulnerability exists. Developers or users may not need to consider applicable vulnerability if vulnerability points do not significantly affect risks of applications in certain situations or conditions, even, if they were found using vulnerability analysis tools. However, a risk evaluation system is also required for vulnerability management systems because security patches or application adjustments are mandatory.

The Common Vulnerability and Exposures (CVE) is a database for cyber security weak points; it manages occurrences of those weak points by years, sequences and other characteristics. Many US government agencies and diverse companies such as Apple, Microsoft and others all participate in CVE; currently, the US Department of Homeland Security manages it.CVE was first created
by MITRE, a nonprofit organization in 1999 and the National Institute of Standards and Technology (NIST)'s coordination helped with the systematic management. When weak points are detected, one can request a registration to MITRE, after which request the CVE management then discusses the candidate and potentially assigns a CVE-ID (FIRST!, 2016).

The CWE/SANS 25 is the list that contains the most dangerous errors of software. MITRE under DOD and Software security experts from the US and Europe made CWE/SANS 25 in order to minimize damages that occur from severe errors of software. This list gets developed by categorizing top 25 vulnerability by assigning scores to vulnerability list provided by CWE using CWSS (Common Weakness Scoring System). The most recent list was released in 2011 and it gets sorted to 3 categories. These 3 categories are incomplete interaction among components such as “SQL injection”, dangerous resource management such as “Buffer Overflow”, porous defense system such as ‘non-certify of important function’ (Christy, 2017). Like this, 25-vulnerability list is the guideline for the safe software development and ways to minimize security damages due to software.

The Common Vulnerability Scoring System (CVSS) is a system to evaluate degree of risks to embedded vulnerabilities. Developed by Carnegie Mellon University and National Institute of Standards and Technology, it is currently managed by the Forum of Incident Response and Security Teams. The first official version was V2.0 presented in 2005; it was updated as V3.0 in June 2015 (Coley, 2016). CVSS has three evaluation items the basic group, time group and environment group. Each evaluation item has detailed evaluation tables to calculate weak points. The basic group evaluates general degrees of risk that are well-known. The time group and environment group are used under special circumstances for detected vulnerabilities from special time flow or other situations.

The Common Weakness Scoring System (CWSS) is the system that evaluates security vulnerabilities that software possesses quantitatively and the US government, academics and industry are all involved in its research. For quantitative evaluation, CWSS calculates scores of 3 sectors separately. The three sectors consist of the base finding metric group (a group that evaluates quantitative scores of the security vulnerability itself), the attack surface metric group (efforts that attackers need to put in during software attacks using the relevant security vulnerability) and the environmental metric group (environmental factors that alter the strength of the relevant security vulnerability). All 3 sectors have 16 detailed items, the midpoints of each sector get calculated and then those midpoints are multiplied to calculate the final CWSS value. The current CWSS is Version 1.0.1 and it was last updated in 2014.

RESULTS AND DISCUSSION

System architecture: Our proposed technique has 5 stages. The first stage is the basic analysis phase which detects vulnerability by using the static analysis tool and the dynamic analysis tool, these two tools minimize problems with the dynamic analysis tool through interaction. The second stage is the analysis stage of the basic vulnerability analysis results; it analyzes results from 1st stage, so that, results can be used during the third and fourth stages. The third stage is the generation of a customized rule set and in-depth test stage which executes an in-depth static analysis using information generated from second stage. Information used at this stage was analyzed from the second stage where it originated from the dynamic analysis module from the first stage. Results from the fourth stage go through in-depth tests after generating exploits using information from the second stage; this also went through the static analysis module of first stage. The relevant application’s degree of risk gets evaluated by collecting results from the fifth stage. The fifth stage reports lists of detected vulnerabilities; the degree of risks is computed using detection accuracy of the vulnerability lists. From the above mentioned stages, results from the first stage static analysis module get analyzed once again at the in-depth test of the fourth stage to minimize over detections and incorrect detections of the static module of the first stage which increases detection accuracy.

Additionally, results from the dynamic analysis module of the first stage go through an in-depth test of the third stage to verify them at the source code. Both static and dynamic analysis modules then get cross checked to increase detection accuracy. This is the new technique that is proposed in this study. Moreover, using information of detected vulnerabilities, a new methodology that evaluates the level of dangers that applications are situated is also proposed. Figure 1 shows the general structure of the proposed intelligent vulnerability analysis technique structure.

Vulnerability analysis process

Basic vulnerability analysis stage: This stage analyzes application’s weak points by using both static module and dynamic module. This stage is constructed as shown in Fig. 2. Source code collection (A) receives source codes of applications and prepare to perform static analysis. At this time, source code collection (A) sends
pre-collected information of source codes to application type analysis (D) of dynamic module to determine the type of application. Rule setting (B) sets rules for vulnerability that are to be analyzed and execute the rule testing at analysis engine (C).

Analysis of rules gets performed using pattern matching engine and symbolic execution engine. Application type analysis (D) of dynamic module uses information received from Source code collection (A) to analyze types such as applet, web application independent systems and others. Then, fuzzing rule for dynamic analysis and build setting for execution get set. Lastly, based on types of application, vulnerability list gets generated through basic dynamic tests using network fuzzing, memory fuzzing and other fuzzings.

Results stage for basic vulnerability analysis: The second stage of the intelligent vulnerability analysis technique, i.e., the analysis stage, analyzes results from the first stage and generates information for the third and fourth stage’s weak-points analysis. Figure 3 illustrates the processes of second stage.

When results from the basic vulnerability weak-points analysis stage are sent to the second stage, it analyzes whether those results are detected from the same locations of the source codes or not. Due to the
Fig. 3: Step 2, Analysis process for basic vulnerability analysis results

characteristics of weak points, some cannot be detected using either static or dynamic analysis methods. Therefore, result categorization (A) develops a list of top 10 of the CWE/SANS 25 that both modules detect commonly and also a list that any one module can detect.

**Customized rule set generation and in-depth test stage:**
A three staged customized rule set generation and In-depth test stage is the static analysis module that goes through in-depth tests to enhance accuracy of results from dynamic module received from the second stage. Figure 4 illustrates this three staged processes.

The vulnerability list (A) that comes to the third stage is the same as the ones that completed dynamic analysis but were not decided as the same causes with static analysis. From rule type analysis, (B) the vulnerability list gets analyzed and categorized as to which rule type it corresponds to. Separate from the existing 10 items, a new empty rule base form (C) gets generated. To re-check the delivered vulnerability lists, customized rules are generated from the new rule generation (D). At this stage in order to find common detection results between results from static analysis and dynamic analysis, detected vulnerability gets categorized. This stage is to find vulnerability list that is detected as common vulnerability from same code location of applications; except common lists, results of static analysis uses dynamic technique, results of dynamic analysis uses static technique to re-analyze vulnerability of application.

**Exploit generation and in-depth test stage:** Exploit generation and in-depth test stage is the fourth stage and it is opposite to the third stage. Vulnerability lists that were either mapped of not mapped from detection results of dynamic analysis module get transferred. These vulnerabilities were detected using the static analysis methods; there is a chance of them being due to false or over detection; it is required to re-check the results using dynamic analysis module to increase the detection accuracy. Figure 5 illustrates the four-stage process.

Exploit process (A) for each vulnerability is defined after analyzing the vulnerability list from second stage. Since, the dynamic module does its analysis during the execution of the application, codes that were detected with vulnerabilities can be executed at this stage. At this time, build information for the applications and application itself get executed first, then information which can reach towards vulnerability code gets delivered (B). After the process is defined, the exploit form is completed by
Step-5
Result list of each step
Apply weighting
Apply modified CWSS to calculate vulnerability score
Final Risk Score Calculation
Application risk estimate by result 1, 2, 3, 4 step with modified CWSS

Fig. 6: Step-5: Process for evaluation of degree of risk

<table>
<thead>
<tr>
<th>Table 1: Detection accuracy of results by stages</th>
<th>Accuracy</th>
<th>Probability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>①</td>
<td>High</td>
<td>81</td>
</tr>
<tr>
<td>②</td>
<td>Low</td>
<td>64</td>
</tr>
<tr>
<td>③</td>
<td>Mid</td>
<td>69</td>
</tr>
<tr>
<td>④</td>
<td>Very high</td>
<td>88</td>
</tr>
</tbody>
</table>

Result ④ is assessed to have high accuracy than result ②. Moreover, results of differences (④-(③∩②)) from 4th stage goes through module analysis to enhance detection accuracy in results of ④. Accuracy that were detected by each stage is illustrated in Table 1. Test bed was structured first and results were drawn. Detection accuracy of each vulnerability is quite similar to ones that human analyzed manually.

Evaluation scores and mid scores of vulnerability self degree of risk: The vulnerability self degree of risk uses portions of the CWSS calculation logic. The CWSS's security vulnerability degree of risk score calculation uses the base finding metric group, attack surface metric group and environmental metric group. The environmental metric group applies external factors such as type of application, detection possibilities of the relevant vulnerability, degree of spread and other external causes. However in this study, the environmental metric group does not affect vulnerability significantly and it is removed from the vulnerability evaluation. Therefore in this study, only the base finding metric group and attack surface metric group are multiplied independently in order to evaluate the degree of risks of the vulnerability. Equation 2 illustrates the scoring methods of the base finding metric group:

\[
\text{Scores of degree of risk of vulnerability} = \text{Base finding} \times \text{Attack surface}
\]

Table 2 shows the abbreviations used in the following equations:
Fig. 7: Details of proposed intelligent vulnerability analysis technique

Table 2: Abbreviations used in formulas

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Full words</th>
</tr>
</thead>
<tbody>
<tr>
<td>TI</td>
<td>Technical Impact</td>
</tr>
<tr>
<td>AL</td>
<td>Acquired privilege Layer</td>
</tr>
<tr>
<td>IC</td>
<td>Internal Control effectiveness</td>
</tr>
<tr>
<td>RL</td>
<td>Required privilege Layer</td>
</tr>
<tr>
<td>SC</td>
<td>Deployment Scope</td>
</tr>
<tr>
<td>AS</td>
<td>Authentication Strength</td>
</tr>
<tr>
<td>AP</td>
<td>Acquired Privilege</td>
</tr>
<tr>
<td>FC</td>
<td>Finding Confidence</td>
</tr>
<tr>
<td>RP</td>
<td>Required Privilege</td>
</tr>
<tr>
<td>AV</td>
<td>Access Vector</td>
</tr>
<tr>
<td>IN</td>
<td>Level of Interaction</td>
</tr>
</tbody>
</table>

Base finding = \{(10\times TI + 5\times (AP + AL) + 5\times FC) \times (TI + IC) \cdot 4.0\}

The weighted score of 10 is used for technical affect due to the relevant vulnerability, weighted scores of 5 are used for reports of successful attacks and reliability about whether the relevant vulnerability actually exists or not is used to compute the scores. A unique characteristic is that the scores for the relevant vulnerability is not computed once the technical affect is decided to be 0. The score for AL is '1'; since, FC is also notified of the vulnerability, score is also '1'. Also, in order not to use IC, a score of '1' gets computed for the computing base finding scores. Attack surface metric score:

ASM score = \{(20\times (RP + RL + AV) + 20\times SC + 15\times IN + 5\times AS) / 100\}

In order to use the relevant vulnerability, the RP, RL, AV and SC are weighted each with 20%, the IN and AS are weighted with 15 and 5%, respectively. RL was not used in this study and we computed it as '1'. The Mid score uses the risk score that was computed from Eq. 1 and the accuracy of the relevant vulnerability. The mid score gets computed as the degree of risk of vulnerability score times the weights for accuracy. Detection accuracy is categorized into four types, based on the methods of the proposed technique. Table 3 shows the weights for accuracy.

Mid score = Degree of risk of vulnerability score \cdot weights for accuracy

Evaluation scores and mid scores of vulnerability self degree of risk: The final evaluation of degree of risk of the application is decided by considering the number of detected vulnerabilities and frequency of occurrences. In order to reflect these factors, the scope of vulnerability self-risk scores is categorized into three sections. The first section is for under 65 indicated with 1–9; The second section is over 66 and under 85 as indicated with 10–90 and the third section is over 85 indicated with 100–200 (Table 4).

Degree of risk presents the number of vulnerability detected at application using relative concept. Empirically risk weight was assigned-single vulnerability with high
Fig. 8: Report of detail results for application vulnerability analysis

Table 4: Ways to describe final vulnerability scores by sections of mid scores

<table>
<thead>
<tr>
<th>Section of mid scores</th>
<th>Scores of degree of risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 65</td>
<td>1st digit: 1–9</td>
</tr>
<tr>
<td>65–84</td>
<td>10th digit: 10–60</td>
</tr>
<tr>
<td>Above 85</td>
<td>100th digit: 100–200</td>
</tr>
</tbody>
</table>

scores of CWSS is more dangerous than much vulnerability with low scores. Namely, degree of risk is the adjusted value to clarify application’s degree of risk intuitively bases on mean scores.

Implementation: In order to test proposed technique, vulnerability analysis system was implemented as prototype. System was made up with both static and dynamic modules; 2 static analysis tools and 7 dynamic analysis tools were used. Table 5 indicates analysis tools for test bed.

The system used a CPU quad core 2.66G Hz, RAM 4 Gb, Windows 7, HD 500 G which was used to install all analysis tools and evaluation page was used in jsp for the reports on each analysis tool’s linkages. Data set for assessment includes 15,277 Java source file known as Juliet code, 2 static analysis technique and 7 dynamic analysis technique. Using them, vulnerability detection test is performed by using and without using proposed technique. The application analysis data was composed of analyzed items, project items to show analyze results, report items to show project results and setting items to set up vulnerability items for detections. Figure 8 is a screen shot of the embodied system’s vulnerability analysis.

The above shows the detailed results after analysis of the particular open source’s vulnerability that is uploaded to the system. Detailed information of the vulnerability such as where errors are found or the vulnerable function’s details can be shown by selecting the relevant vulnerability.

Performance evaluation: Table 6 presents the number of total detections, proper detections and false detections of the existing static analysis tools (PMD, Find Bugs) and proposed technique for top 10 items of CWE/SANS 25. Excluding only one item, the proposed technique presents many proper detections from all items, showing that its capabilities for detecting vulnerability is superior over existing static analysis tools.
Table 6: Comparison of analysis results of existing static analysis tool and static and dynamic mixed analysis results that apply proposed technique

<table>
<thead>
<tr>
<th>Top 10 of SANS 25</th>
<th>PMD</th>
<th>Find bug</th>
<th>Apply proposed method (PMD + peachpuzzer)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Proper</td>
<td>False detect</td>
</tr>
<tr>
<td>SQL input</td>
<td>482</td>
<td>299</td>
<td>183</td>
</tr>
<tr>
<td>OS command input</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Buffer overflow</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cross site script</td>
<td>61</td>
<td>45</td>
<td>16</td>
</tr>
<tr>
<td>Allowance of critical functions without proper authorization</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inadequate authority</td>
<td>14</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Hard coded critical information</td>
<td>112</td>
<td>107</td>
<td>5</td>
</tr>
<tr>
<td>Clear text transmission and save of critical information</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>File uploads</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Inappropriate input value used in security function decisions</td>
<td>16</td>
<td>11</td>
<td>5</td>
</tr>
</tbody>
</table>

Because the result of mix usage of PMD and peachpuzzer was the most superior over all other mix of techniques; results of mix usage of PMD and peachpuzzer is used in this research.

CONCLUSION

In order to increase accuracy of detecting vulnerabilities that are embedded into the application, this study presented a vulnerability detection technique which increases accuracy of results. A new technique was proposed, i.e., the customized static analysis rule gets generated and an intelligent exploit code is generated to re-verify both static and dynamic modules to increase the accuracy of vulnerability detection. Moreover, new scores for application risk evaluation using self-risk scores of vulnerability are provided by CWSS and proposed detection accuracy is also suggested.

We found that the vulnerability detection capabilities of our proposed technique have better results than the capabilities of existing static and dynamic analysis tools. Moreover, the risk evaluation scores of applications using the proposed methods were suggested. This relevant application shows level of dangers at a look this proposed technique has strong points over existing analysis methods.

RECOMMENDATIONS

For the future research, environmental matrix calculation which was not used in this study will be added to find out its impact, its spread, its attack-ability and others this will result in full details of category and to improve reliability of weak degree of risk calculations.

REFERENCES


Christey, S., 2011. CWE/SANS top 25 most dangerous software errors. MSc Thesis, MITRE Institute, Bedford, Massachusetts.


