STATIC CODE ANALYSIS TOOLS: A SYSTEMATIC LITERATURE REVIEW

Darko Stefanović, Danilo Nikolić, Dušanka Dakić, Ivana Spasojević, Sonja Ristić

Abstract

Static code analysis tools are being increasingly used to improve code quality. Such tools can statically analyze the code to find bugs, security vulnerabilities, security spots, duplications, and code smell. The quality of the source code is a key factor in any software product and requires constant inspection and supervision. Static code analysis is a valid way to infer the behavior of a program without executing it. Many tools allow static analysis in different frameworks, for different programming languages, and for detecting different defects in the source code. Still, a small number of tools provide support for domain-specific languages. This paper aims to present a systematic literature review focusing on the most frequently used static code analysis tools and on classifying the presented tools according to the supported both general-purpose and domain-specific programming languages and the types of defects a specific tool can detect.

Keywords: static code analysis; tools; defects; reengineering; literature review;

1. Introduction

At the Faculty of Technical Sciences, various forms of application of information technologies are studied, such as [1] [2] [3] [4] [5], in addition, information technologies find application in areas such as [6] [7], in this paper, the application of information technologies and tools to static code analysis will be described.

Since the beginning of software development as a scientific discipline, software developers have lacked a method for source code quality assessment. Most software developers rely on the compiler and the fact that the software product works successfully until an error is caused by poor code quality or a bug that disrupts the operation of the software solution occurs.

The quality of the source code is a key factor in any software product and requires constant verification and monitoring. Static analysis is used to maintain and improve the quality of the source code. Static analysis of program code has been used since the early 1960s to optimize the operation of compilers [8]. Later, it proved useful for debugging tools, as well as for software development frameworks. There is a growing number of tools that allow static code analysis, many of which are open-source tools and allow an analysis of several different programming languages [9].
Static analysis tools are used to generate reports and to point out certain deviations from the prescribed code quality standards. However, these tools do not allow automatic modifications to the source code. The decision to change the way the previously written code is structured remains in the hands of software developers. Static code analysis tool helps software developers by, in addition to generating a report, stating the cause of a particular defect, as well as how this defect can be corrected.

However, the question remains on how to approach the defects, to eliminate the highlighted shortcomings, and reengineer the source code.

When it comes to static code analysis tools, a big problem is a fact that there are a large number of tools that provide this type of analysis. The number of tools is constantly increasing, and it is necessary to classify existing tools according to the programming language for which they have the support and the types of defects they detect.

Static code analysis is performed on the source code of the program and therefore the programming languages in which the program is written are very important. From the side of static code analysis, it is necessary to separate the types of programming languages into general-purpose programming languages (GPL) and domain-specific programming languages (DSL). In this literature review, there is a classification of tools that provide support for these two groups of programming languages.

Studies that are included in the literature review are set for a certain period of time. This paper reviews published studies in the previous ten years, conducted based on the guidelines for performing systematic literature reviews in software engineering by Barbara Kitchenham [10], with a focus on displaying the most commonly used tools divided by the programming languages for which they have support and defects they detect.

The remainder of the paper is organized as follows. Section 2 presents the basic concepts of static code analysis tools. Section 3 describes the systematic literature review methodology, presents conducted planning the review phase of the systematic literature review, and presents conducting the review phase of the systematic literature review, where primary studies were selected and the review results presented. Section 4 discusses the systematic literature review results in detail, and Section 5 concludes the paper and suggests future research.

This literature review will be used as a basis for conducting further experiments on static code analysis tools.

2. Theoretical foundations

This section will present what manual code review is, how it is performed, how static code analysis and its tools help this process, and what are domain-specific and general-purpose programming languages.

2.1. Manual code review

Code review is used to improve the maintenance process of a software product, increasing its reliability and security. Code review activity can reduce the number of required updates in the source code, and can be easily integrated into the software development lifecycle so that more experienced software developers do each code review. However, a manual review of the source code takes a very long time [11]. To conduct an efficient manual code review, reviewers must know all possible defects and inaccuracies before carefully checking all source code. As a substitute for a lengthy manual source code review, the use of a static code analysis tool is recommended. These tools are faster and contain the knowledge to perform the scan. However, these tools cannot completely replace the manual review, but they facilitate this process.

2.2. Static code analysis

The static code analysis process is useful not only for optimizing the operation of the compiler (which was the original purpose) but also for detecting irregularities and possible defects. In this way, it is possible to create tools that will help developers to understand the behaviour of a program and to identify various defects in a program without its execution. The tools used for static code analysis are programs that explain the behaviour of other programs [8].

Static code analysis is significantly faster than conventional testing and can detect any defect visible in the program's source code. If the tools for static code analysis are compared with the manual code review by software developers, it can be concluded that the code review using the tool is also much faster and more efficient [11]. However, for static code analysis to detect any defect, it must be visible in the source code.

In addition to the amount of context required to identify a defect, many defects can only be found in a particular representation of the program. Figure 1 shows a matrix formed according to the type and visibility of the defect. High-level problems are often only visible in program design, while implementation errors can often only be spotted by examining program source code. Object-oriented languages such as Java have a large number of libraries, which makes it easier to understand the design by examining the source code.
Static code analysis tools look for a specific set of patterns or rules in the source code, very similar to the way antivirus programs detect viruses. More advanced tools allow practitioners to add new rules to a set of predefined ones. These tools cannot find an error if patterns or rules do not specify such behaviour. The most significant value of a static code analysis tool is the ability to identify common programming defects.

The standard code review cycle includes four main phases [11]:
1. establish goals,
2. run the static analysis tool,
3. review code (using the output of the tool),
4. make fixes.

In addition to the standard cycle, Figure 2 shows several potential feedbacks or minor iterations between standard cycle steps, making the cycle more complex than the standard procedure.

![Fig. 2. Code review cycle [11]](image-url)
2.3. General-purpose (GPL) and domain-specific programming languages (DSL)

A GPL is a computer language that is broadly applicable across application domains and lacks specialized features for a particular domain. Static analysis tools and quality inspection tools, in general, are prepared to analyse the source code of GPLs.

Contrary, a DSL is specialized in a particular application domain. The line is not always sharp, as a language may have specialized features for a particular domain but be applicable more broadly, or conversely, may be capable of broad application but in practice is used primarily for a specific domain [12]. Professionals in many different areas can develop and describe problem solutions that take the shape of source code artefacts using DSLs that are specially defined for their field [13].

Static code analysis is a great challenge when it comes to DSLs, so this paper will highlight the tools that can contribute to static analysis in this group of programming languages.

2.4. A Taxonomy of Software Security Errors

In order to successfully classify all defects presented in the studies included in this literature review, it is necessary to introduce a unique taxonomy according to which all presented defects will be classified.

A simple, intuitive taxonomy of software security is proposed in [14], called Seven Pernicious Kingdoms. This classification scheme is amenable to automatic identification using static analysis tools for detecting real-world security vulnerabilities in software. In this way, this taxonomy encompasses all security defects that static code analysis can detect and for this reason, this taxonomy is used in research.

The classification divides the defects into seven (plus one) high-level kingdoms, seven of these kingdoms are dedicated to errors in source code, and one is related to configuration and environment issues. It is presented in order of importance to software security [14]:

1. input validation and representation,
2. API abuse,
3. security features,
4. time and state,
5. errors,
6. code quality,
7. encapsulation,
8. environment.

All studies included in this literature review use different taxonomies to classify the defects that the presented tools detect. For this reason, a uniform taxonomy is used in this paper and all defects were classified according to this presented taxonomy.

3. Methodology

To conduct this systematic literature review, a procedure for systematic reviews developed by Barbara Kitchenham [10] is followed. According to Kitchenham, a systematic literature review can be summarized into three main phases: Planning the Review, Conducting the Review, and Reporting the Review.

3.1. Planning the Review

The first activity of Planning the Review phase should be an elaboration of the need for a systematic literature review, by reviewing existing literature reviews of the subject [10]. However, there are no explicit, systematic literature reviews on the subject of static code analysis tools. There is only threat analysis in software systems in general [15]. Other sources of information about static code analysis tools could be the related work chapters of studies that are presenting a novelty on the subject. However, these studies only mention previous research in the field of static code analysis relevant to their particular topic of interest. A systematic review of the literature presents the static code analysis tools most commonly used in studies and the supported programming languages. Studies describing a static code analysis tool or tools and analysing their defect detection in one or more programming languages have been searched and included in the literature review.

Based on the guidelines from Kitchenham [10], the following research questions are formulated:

RQ1: Are there static code analysis tools applicable to any GPL?
RQ2: Is there a static code analysis tool that applies to a DSL?
RQ3: Which tools are most commonly mentioned in the primary studies?
RQ4: Is there a tool that detects all types of defects?

For this literature review, the following databases were searched:

1. SCOPUS and
2. Google scholar. 

Search terms defined for search in these databases is: “Static code analysis” and “tools” and “detection” and 
pubyear>=2010. The search resulted in 346 papers. 

Inclusion criteria applied to papers are: 
IC1: Papers have to be written in English. 
IC2: Paper has to present tools based on static code analysis. 
IC3: Paper has to present supported programming languages by static code analysis tool(s). 

Exclusion criteria applied to papers are: 
EC1: Duplicate papers should be removed. 
EC2: If one author had more than one paper regarding the same tool, only one paper should be included in the review. 

For this literature review, the data extraction strategy is developed. For each primary study, the following features will 
be extracted in order to answer the research questions: 
1. publication year and source type, 
2. proposed tool(s), 
3. type(s) of defect(s) they detect, 
4. supported programming languages (DSL/GPL). 

3.2. Conducting the Review 

After applying defined inclusion and exclusion criteria on available papers, papers were critically appraised based on their 
relevancy and type of information they contained. Finally, there were 22 primary studies available for data extraction. 

The flow diagram of the systematic literature review process is presented in Fig. 3. 

![Systematic literature review flow diagram](image)

Search through data sources with previously defined search terms resulted in the identification of 346 primary studies. 
Based on the title and abstract, 260 papers were excluded from further research. A review based on inclusion and exclusion 
criteria is conducted on the paper and resulted in the inclusion of 27 primary studies. The review based on a critical 
assessment of the study resulted in the exclusion of 5 studies from further research. Finally, 22 primary studies are 
included in the literature review. 

3.3. Data extraction and summarization 

In this chapter, the main characteristics of primary studies are summarized and presented, following with the tables that 
present extracted data. Interpretation and discussion of the presented results will be described in the next chapter. 

Fig. 4 presents primary studies that present static code analysis tools by year of publication. The line diagram shows 
that two peaks in the number of published studies occurred, one in 2012 and one in 2016.
Table 1. presents source types of primary studies, showing that 37% of the papers were published as journal articles, 59% of conference proceedings, and one master thesis (4%).

<table>
<thead>
<tr>
<th>Source type</th>
<th>Studies</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal article</td>
<td>[16], [17], [18], [19], [20], [21], [22], [13]</td>
<td>37%</td>
</tr>
<tr>
<td>Conference proceedings</td>
<td>[23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35]</td>
<td>59%</td>
</tr>
<tr>
<td>Master thesis</td>
<td>[36]</td>
<td>4%</td>
</tr>
</tbody>
</table>

Table 1. Source types of primary studies

In the review protocol, three research questions were formulated. Figure 5 presents extracted data that answers the first research question-RQ1: Are there static code analysis tools applicable to any GPL?

Some of the presented static code analysis tools have support for several GPLs, while other tools have support for only one programming language. Most static code analysis tools have support for those programming languages that are most commonly used in the industry. The reason for this is that in this way, the creators of the static code analysis tool provide a higher number of users of their tool.

Fig. 5. Supported programming languages by the tool(s) presented in papers

Table 2. presents extracted data that answers the second research question-RQ2: Is there a static code analysis tool that applies to a DSL?

Within the selected studies in the literature review, only three studies (14%) represent tools and support for DSLs. In contrast, other tools (86%) in the studies analyze the accuracy and efficiency of static code analysis tools that have support for one or more GPLs.
Figure 6 presents extracted data that answer the third research question-RQ3: Which tools are most commonly mentioned in the primary studies?

Six tools were presented, which were most often analysed and presented in primary studies. These are tools that have been presented three or more times in different researches. In comparison, other tools presented within the studies included in this literature review were analysed in two or one studies.

![Number of presentations within the study](image)

Fig. 6. The most common static code analysis tools presented in primary studies

In Table 3 are given the extracted data that answer the research question-RQ4: Is there a tool that detects all types of defects?

The problem observed in the extraction of data related to the types of defects to be detected is that the studies do not clearly and precisely highlight the types of defects that the proposed tools detect. Some studies do not clearly (or at all) highlight the types of defects that the proposed tools detect, and such studies [13] [26] [29] [30] [33] are excluded from this classification by type of defect.

Another problem observed is that studies use different taxonomies to classify defects. In this research, a unique taxonomy has been established, which is called Seven pernicious kingdoms [14], and all types of defects that various studies point out have been classified according to it.

Table 3 shows the classification of studies according to the types of defects detected within the presented tools.

<table>
<thead>
<tr>
<th>Seven Pernicious Kingdoms</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input validation and representation</td>
<td>[16], [17], [18], [19], [20], [21], [23], [24], [25], [28], [32], [34], [35], [36]</td>
</tr>
<tr>
<td>API abuse</td>
<td>[20], [28], [31], [36]</td>
</tr>
<tr>
<td>Security features</td>
<td>[16], [20], [21], [25], [27], [28], [36]</td>
</tr>
<tr>
<td>Time and state</td>
<td>[20], [21], [31], [32]</td>
</tr>
<tr>
<td>Errors</td>
<td>[16], [19], [20], [21], [23], [28], [32]</td>
</tr>
<tr>
<td>Code quality</td>
<td>[16], [20], [21], [23], [28], [31], [32], [35]</td>
</tr>
<tr>
<td>Encapsulation</td>
<td>[20]</td>
</tr>
<tr>
<td>Environment</td>
<td>[28]</td>
</tr>
</tbody>
</table>

Table 3. Primary studies classified by type of defects that presented tools detect

4. Discussion
The results are commented in this section of the conducted systematic literature review, presented in section 3.3.

The highest number of the reviewed studies were published in the years 2012 and 2016. However, the number of papers on an annual basis is uniform, and research related to static code analysis tools is becoming more common, and more importance is attached to these tools in order to increase the quality of the source code of the software solution.

Table 2. Primary studies classified on DSL and GPL

<table>
<thead>
<tr>
<th>Programming language</th>
<th>Studies</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSL</td>
<td>[13], [29], [35]</td>
<td>14%</td>
</tr>
<tr>
<td>GPL</td>
<td>[16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [30], [31], [32], [33], [34], [36]</td>
<td>86%</td>
</tr>
</tbody>
</table>
Special mention should be made of the Master thesis derived from Alto University, School of Science, as a study that describes in the most detail the presented tools and the experiment on those tools, which makes a great contribution in this area.

In order to answer the research question RQ1, the tools presented in the studies were analyzed, it is concluded that each of the proposed tools has support for at least one GPL. Some of the tools, such as the tool presented in [33], have support for a number of GPLs, while the tool presented in [22] has support for just a single GPL. In the context of the question RQ1 it is important to consider the answer to the question RQ2. Lessons learned from the development of tools that can support DSL can be applied to the development of tools for the GPL.

The second research question RQ2 is answered positively. Creating a static code analysis tool that will provide support for DSLs is of great importance in this area. Of the studies presented in this literature review, three studies highlight tools for static code analysis of programs written by means of DSLs. The framework within the study [35] is presented, and the way of using and efficiency of using the tool over three static code analysis of programs written by means of DSL is explained. On the other hand, the tool, which is otherwise intended for GPL, is presented in the study [13] as a possible solution for conducting static code analysis over DSLs. In this study, they highlight a plug-in that defines source code quality control rules for any DSL. Also, within [13], an experiment is performed on a DSL, and the obtained results were presented. Based on the presented, it is concluded that within the already developed tool for static code analysis, there is a way to provide code overview and detect defects over any language specific for the domain.

To answer the research question RQ3, the tools presented in the studies were analyzed. It is concluded that some tools for static analysis are more often used in research than others. It mainly depends on which GPLs the proposed tools support. Thus, the Cppcheck tool is used in five different studies [16] [21] [28] [32] [31], FindBugs in four [17] [21] [26] [31], as well as Splint studies [16] [21] [28] [32], followed by the SonarQube tool presented in three studies [13][17][33], as well as the PMD tool studies [16] [21] [26] and the static code analysis tool Flawfinder is also presented in three studies [16] [21] [28]. The reason for this lies in the fact that these tools are free and are most often used for static code analysis, so their accuracy and metrics offered within the programming languages for which they are supported are often examined.

The research question RQ4 is answered in the negative, i.e., no tool or study includes all types of defects according to the presented taxonomy. The study [20] presents seven of the eight main types of defects identified by the classification scheme. Other studies included in this review detect only certain types of defects. The biggest problem in distinguishing types of defects from tools and studies and their classification is that most studies use different schemes for defect classification.

5. Conclusion

Static code analysis tools are increasingly finding their application and use in the work of software developers in different positions. New tools are being developed, and existing tools are being improved every day, offering their customers an increasingly high-quality service by creating confidence in the results of static analysis and enabling them to avoid certain code defects and make their software solution better.

Still, static code analysis tools allow the analysis of only a certain number of programming languages (RQ1). However, the range of programming languages for which they have support is continuously expanding. In comparison, it is expected that in some future period, it will work on creating and improving several tools that will have support for DSLs (RQ2). For now, only a few tools have support for DSLs with the specification of rules embedded in the tool via a plug-in.

One of the research questions in this paper is which tools are most often used in research (RQ3). The results show that certain tools are used more often in studies than others. It is concluded that the research is focused mainly on tools that are more often used in practice and that have support for certain GPL. In this way, the research compares the results of the accuracy and efficiency of the presented tools.

Based on the obtained results, we provide an answer to the research question RQ4, it is concluded that no tool can detect all defects defined by the presented classification, while a combination of certain tools can solve the presented problem. Also, the same conclusion is used in the paper for all studies, which leads to an easier understanding of the difference between the tools presented in the studies.

Further research involves performing experiments on the presented tools to compare their advantages and disadvantages in terms of static code analysis, as well as expanding the scope of studies and tools for static code analysis and their classification according to the supported programming languages and the defects they detect.

The use of tools for static code analysis and continuous work on improving and maintaining the quality of source code provides higher quality software solutions and represents a step forward in software development.

6. References


Kitchenham, B. (2004) Procedures for Undertaking Systematic Reviews, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1)


