

Optimizing Symbolic Execution Through Taint Analysis and Path Prioritization

Bachelor thesis

Natural Science Faculty of the University of Basel
Department of Mathematics and Computer Science
Databases and Information Systems (DBIS) Group
<https://dbis.dmi.unibas.ch/>

Examiner: Dr. Marco Vogt
Supervisor: Prof. Dr. Christopher Scherb

Ruben Hutter
ruben.hutter@unibas.ch
2020-065-934

02.07.2025

Acknowledgments

I would like to express my sincere gratitude to Prof. Dr. Christopher Scherb for his supervision and guidance throughout this thesis. His expertise in program analysis and symbolic execution provided essential direction for this research.

I am grateful to Dr. Marco Vogt for his role as examiner and for facilitating the opportunity to pursue this research topic. His feedback and suggestions helped refine both the technical approach and the presentation of this thesis.

I would like to thank Ivan Giangreco for providing the LaTeX thesis template used for this document, which greatly facilitated its formatting and structure.

I also acknowledge my fellow student Nico Bachmann for developing the Schnauzer visualization library, which enhanced the presentation and analysis of the results.

I thank my family and friends for their encouragement and support during my studies, which made completing this thesis possible.

Finally, I acknowledge the developers of the angr binary analysis framework, whose comprehensive platform enabled the implementation of the techniques described herein.

Abstract

Symbolic execution is a powerful program analysis technique widely used for vulnerability discovery and test case generation. However, its practical application is often hampered by scalability issues, primarily due to the "path explosion problem" where the number of possible execution paths grows exponentially with program complexity. This thesis addresses this fundamental challenge by proposing an optimized approach to symbolic execution that integrates taint analysis and path prioritization.

The core contribution is a novel exploration strategy that moves away from uniform path exploration towards targeted analysis of security-critical program behaviors. The approach prioritizes execution paths originating from memory allocations and user input processing points, as these represent common sources of vulnerabilities. By leveraging dynamic taint analysis, the system identifies and tracks data flow from these critical sources, enabling the symbolic execution engine to focus computational resources on paths influenced by tainted data while deprioritizing paths with no dependency on external inputs.

The implementation integrates this taint-guided exploration strategy with the angr symbolic execution framework, introducing a scoring mechanism that dynamically adjusts path prioritization based on taint propagation. The effectiveness of this optimization is evaluated through comparative analysis, examining runtime efficiency, path coverage quality, and vulnerability discovery capabilities. Results demonstrate that this approach can significantly reduce the search space while maintaining or improving the detection of security-relevant program behaviors, making symbolic execution more practical for large and complex software systems.

Table of Contents

Acknowledgments	ii
Abstract	iii
1 Introduction	1
2 Background	4
2.1 Symbolic Execution	4
2.2 Program Vulnerability Analysis	5
2.3 Taint Analysis	5
2.4 Control Flow Analysis	5
2.5 Angr Framework	6
3 Related Work	7
3.1 Optimization Approaches	7
3.1.1 State Space Reduction	7
3.1.2 Performance and Compositional Analysis	7
3.2 Integration Approaches	8
3.3 Security-Focused Targeting and Research Gap	8
4 Taint-Guided Exploration	9
4.1 Core Approach	9
4.2 Taint Source Recognition	10
4.3 Dynamic Taint Tracking	10
4.4 Path Prioritization	11
4.4.1 Adaptive State Pool Management	12
4.5 Exploration Depth Control and Vulnerability Probability	13
5 Implementation	14
5.1 Tool Architecture and Workflow	14
5.1.1 Core Components	14
5.1.2 Workflow Overview	15
5.2 Core Implementation Details	16
5.2.1 TraceGuard Class Architecture	16
5.2.2 Function Hooking Strategy	16

5.2.3	Taint Detection and Propagation	17
5.2.4	Custom Exploration Technique	17
5.3	Architecture Support and Configuration	18
5.3.1	Multi-Architecture Implementation	18
5.3.2	Meta File Integration	18
5.4	Usage and Configuration	19
5.4.1	Command-Line Interface	19
5.4.2	Usage Examples	20
5.4.3	Integration Workflow	20
5.5	Practical Example	20
5.5.1	TraceGuard Execution	20
5.5.2	Analysis Results	21
5.5.3	Interactive Visualization	21
6	Evaluation	22
6.1	Experimental Design	22
6.1.1	Research Questions	22
6.1.2	Evaluation Metrics	22
6.1.3	Experimental Environment	23
6.2	Benchmark Programs	23
6.2.1	Synthetic Test Suite	23
6.3	Experimental Results	24
6.3.1	Vulnerability Detection Effectiveness	24
6.3.2	Execution Time Performance	24
6.3.3	Coverage Analysis and Path Efficiency	25
6.3.4	Statistical Significance and Reliability	26
6.4	Discussion and Analysis	26
6.4.1	Limitations and Practical Considerations	26
7	Conclusion	27
8	Future Work	28
8.1	Enhanced Configuration and Usability	28
8.2	Architecture and Platform Extensions	28
8.3	Scalability and Real-World Applications	29
8.4	Input Source Expansion	29
8.5	Performance Optimization	29
9	Usage of AI	30
	Bibliography	31
	Appendix A Appendix	32
A.1	State Explosion Test Program Source Code	32

1

Introduction

In today's interconnected digital landscape, software security has become a critical concern as applications handle increasingly sensitive data and operate in hostile environments. The discovery of security vulnerabilities before deployment is essential to prevent exploitation by malicious actors, yet traditional testing approaches often fail to comprehensively explore all possible execution scenarios, leaving potential vulnerabilities undiscovered.

Among program analysis techniques, symbolic execution has emerged as a particularly powerful approach for automated vulnerability discovery. Unlike traditional testing that executes programs with concrete input values, symbolic execution treats inputs as mathematical symbols and tracks how these symbols propagate through program computations. When encountering conditional branches, the symbolic execution engine explores multiple possible paths simultaneously, building a comprehensive map of program behaviors. This systematic exploration capability makes symbolic execution especially valuable for security analysis, as it can automatically generate test cases that reach deep program states and trigger complex vulnerabilities such as buffer overflows, integer overflows, and format string bugs.

Challenges in Symbolic Execution. Despite its theoretical power, symbolic execution faces a fundamental scalability challenge known as the path explosion problem. As program complexity increases, the number of possible execution paths grows exponentially, quickly overwhelming computational resources and rendering the analysis intractable for real-world software systems. Modern applications can generate millions of execution paths from relatively small input variations, making exhaustive analysis computationally prohibitive.

The path explosion problem is exacerbated by current symbolic execution engines that typically employ uniform exploration strategies, treating all program paths with equal priority regardless of their potential security relevance, as Figure 1.1 illustrates. This approach fails to recognize that paths processing user-controlled data are significantly more likely to contain vulnerabilities than paths handling only internal program state. Consequently, significant computational resources are often spent analyzing auxiliary program logic while security-critical paths that process external inputs receive no special attention. Consider, for example, a network service that accepts client connections, reads incoming data via

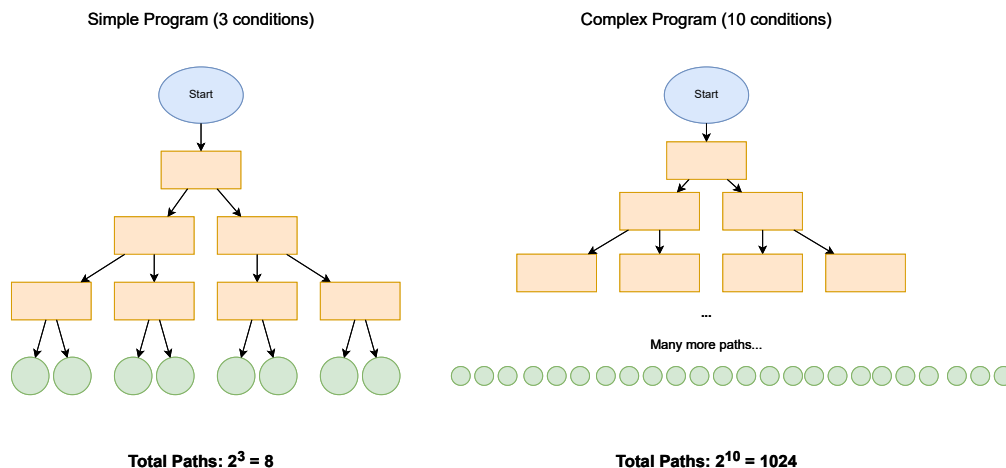


Figure 1.1: Illustration of the path explosion problem: As program complexity increases from 3 to 10 conditions, the number of possible execution paths grows exponentially from 8 to 1024 paths.

network sockets, performs input validation through multiple parsing layers, and eventually stores results using memory copy operations. Traditional symbolic execution would explore all execution paths with equal priority, including those that handle only internal configuration data or administrative functions that never process user input. A security-focused approach should recognize that paths flowing from network input through data processing to memory operations deserve higher priority due to their potential for buffer overflows, injection attacks, and other input-related vulnerabilities.

Thesis Overview. This thesis presents TraceGuard¹, an approach that integrates taint analysis with symbolic execution to enable intelligent path prioritization. Our methodology identifies and tracks data flow from critical sources such as user inputs, guiding the symbolic execution engine to focus computational resources on paths most likely to exhibit security-relevant behaviors. The key insight driving this approach is that not all execution paths are equally valuable for security analysis—paths that interact with user-controlled data are significantly more likely to harbor vulnerabilities than those processing only internal program state. TraceGuard operationalizes this insight through a dynamic taint scoring mechanism that quantifies the security relevance of each symbolic execution state. By prioritizing states with higher taint scores, the symbolic execution engine directs its computational resources toward program regions most likely to contain security vulnerabilities, fundamentally transforming symbolic execution from an exhaustive search into a guided exploration strategy. The main contributions of this thesis are:

- **Taint-Guided Path Prioritization:** An integration of dynamic taint analysis with symbolic execution that uses taint propagation patterns to intelligently prioritize exploration of security-relevant execution paths.
- **Custom Angr Exploration Technique:** Implementation of TaintGuidedExploration, a specialized exploration strategy that extends Angr’s

¹ <https://github.com/ruben-hutter/TraceGuard>

symbolic execution capabilities with security-focused path prioritization.

- **Function-Level Taint Tracking:** A comprehensive taint tracking system that monitors input functions, tracks taint propagation through function calls, and maintains detailed taint information throughout program execution.
- **Adaptive Scoring Algorithm:** A scoring mechanism that dynamically adjusts path priorities based on real-time taint analysis results, enabling the symbolic execution engine to focus computational resources on the most promising program regions.
- **Intelligent Function Hooking System:** A sophisticated hooking mechanism that intercepts function calls to analyze parameter taint status, allowing selective execution of only security-relevant code paths.
- **Practical Implementation and Validation:** A complete implementation using the angr symbolic execution framework, with comprehensive testing demonstrating the effectiveness of taint-guided exploration.

The effectiveness of this optimization is evaluated through controlled experiments comparing TraceGuard against standard Angr symbolic execution techniques using custom-designed test programs with known taint flow patterns. The evaluation examines key metrics including execution time, function call efficiency, and vulnerability detection reliability, demonstrating significant improvements in analysis efficiency while maintaining comprehensive vulnerability detection capabilities. This thesis focuses on binary program analysis using the angr symbolic execution framework, targeting user-space applications written in C/C++ and compiled for x86-64 architectures. The evaluation methodology centers on custom-designed test programs that demonstrate clear taint flow patterns, specifically crafted to evaluate TraceGuard's ability to distinguish between tainted and untainted execution paths. The thesis is organized as follows:

- **Chapter 2** provides essential background on symbolic execution, taint analysis, and the angr framework.
- **Chapter 3** surveys related work in symbolic execution optimization and taint analysis techniques.
- **Chapter 4** presents the conceptual framework and theoretical algorithms underlying TraceGuard's taint-guided exploration strategy.
- **Chapter 5** details the practical implementation, including integration with angr and the design of the scoring mechanism.
- **Chapter 6** presents a comprehensive evaluation comparing TraceGuard's performance against standard symbolic execution techniques.
- **Chapter 7** concludes with a summary of contributions and research implications.
- **Chapter 8** explores potential extensions and future research directions.

2

Background

This chapter establishes the theoretical foundations necessary for understanding the taint-guided symbolic execution optimization presented in this thesis. We examine symbolic execution, program vulnerability analysis, taint analysis, control flow analysis, and the Angr² framework.

2.1 Symbolic Execution

Symbolic execution is a program analysis technique that explores execution paths by using symbolic variables instead of concrete inputs. The program state consists of symbolic variables, path constraints, and a program counter. When execution encounters a conditional branch, the engine explores both branches by adding appropriate constraints to the path condition.

A fundamental challenge in symbolic execution is the path explosion problem. As program complexity increases, the number of possible execution paths grows exponentially, making exhaustive exploration computationally intractable. This scalability issue particularly affects real-world applications with complex control flow structures and deep function call hierarchies. Research has shown that symbolic execution tools designed to optimize statement coverage often fail to cover potentially vulnerable code due to complex system interactions and scalability issues of constraint solvers [6].

Traditional symbolic execution typically employs a forward approach, starting from the program's entry point and exploring paths toward potential targets. However, this method may struggle to reach deeply nested functions or specific program locations of interest. Backward symbolic execution, conversely, begins from target locations and works backwards to identify input conditions that can reach those targets. Compositional approaches combine both techniques by analyzing individual functions in isolation and then reasoning about their interactions.

² <https://angr.io/>

2.2 Program Vulnerability Analysis

Software vulnerabilities represent flaws in program logic or implementation that can be exploited by malicious actors to compromise system security. Understanding these vulnerabilities is crucial for developing effective analysis techniques that can detect them before deployment.

Traditional testing approaches often fail to discover these vulnerabilities because they typically occur only under specific input conditions that are unlikely to be encountered through random testing. Static analysis can identify potential vulnerabilities but often produces high false positive rates due to conservative approximations required for soundness. Dynamic analysis provides precise information about actual program execution but is limited to the specific inputs and execution paths exercised during testing.

Symbolic execution addresses these limitations by systematically exploring multiple execution paths and generating inputs that trigger different program behaviors. However, the path explosion problem means that uniform exploration strategies may spend significant computational resources on paths that are unlikely to contain security vulnerabilities. This motivates the development of security-focused analysis techniques that prioritize exploration of paths involving user-controlled data, as these represent the primary attack vectors for most software vulnerabilities.

2.3 Taint Analysis

Taint analysis tracks the propagation of data derived from untrusted sources throughout program execution. Data originating from designated sources (such as user input functions like `fgets`, `gets`, `read`, or `scanf`) is marked as tainted. The analysis tracks how this tainted data flows through assignments, function calls, and other operations. When tainted data reaches a security-sensitive sink (such as buffer operations or system calls), the analysis flags a potential vulnerability.

The propagation rules define how taint spreads through different operations: assignments involving tainted values result in tainted variables, arithmetic operations with tainted operands typically produce tainted results, and function calls with tainted arguments may result in tainted return values depending on the function's semantics. Dynamic taint analysis performs tracking during program execution, providing precise information about actual data flows while considering specific calling contexts and program states, resulting in reduced false positives compared to static analysis approaches.

2.4 Control Flow Analysis

Control flow analysis constructs and analyzes control flow graphs (CFGs) representing program structure. CFG nodes correspond to basic blocks of sequential instructions and edges represent possible control transfers between blocks. This representation enables systematic analysis of program behavior and reachability properties.

Static analysis constructs CFGs by examining program code without execution, analyzing structure and control flow based solely on the source code or binary representation.

This approach offers comprehensive coverage and efficiency, enabling examination of all statically determinable program paths without requiring specific input values. However, static analysis faces limitations including difficulty with indirect call resolution and potential false positives due to conservative approximations required for soundness.

Dynamic analysis executes the program and collects runtime information, providing precise information about actual program behavior and complete execution context. This approach eliminates many false positives inherent in static analysis and validates that control flow relationships are actually exercised under realistic conditions. However, dynamic analysis results depend heavily on input quality and coverage.

A Call Graph represents function call relationships within a program, where each node corresponds to a function and each directed edge represents a call relationship. Call graphs serve important purposes including program understanding, entry point identification, reachability analysis, and complexity assessment. Call graphs prove valuable for path prioritization strategies, enabling identification of functions reachable from tainted input sources and assessment of their relative importance in program execution flow.

2.5 Angr Framework

Angr is an open-source binary analysis platform providing comprehensive capabilities for static and dynamic program analysis [7]. The platform supports multiple architectures and provides a Python-based interface for research and education [8]. Key components include the *Project* object representing the binary under analysis with access to contents, symbols, and analysis capabilities; the *Knowledge Base* storing information gathered during analysis including function definitions and control flow graphs; the *Simulation Manager* handling multiple program states during symbolic execution and managing state transitions; and the *Solver Engine* interfacing with constraint solvers to determine path feasibility and solve for concrete input values.

Angr supports both static (CFGFast) and dynamic (CFGEmulated) CFG construction. Static analysis provides efficiency but may miss indirect calls, while dynamic analysis offers completeness at higher computational cost. The framework represents program states with register values, memory contents, path constraints, and execution history, providing APIs for state manipulation and exploration control through step functions and various exploration strategies including depth-first search, breadth-first search, and custom heuristics.

The framework’s extensible architecture enables integration of custom analysis techniques, making it particularly suitable for implementing novel symbolic execution optimizations. The symbolic execution landscape includes numerous frameworks targeting different domains and applications, ranging from language-specific tools like KLEE³ for LLVM⁴ bit-code to specialized platforms for smart contract analysis. Angr’s comprehensive binary analysis capabilities, multi-architecture support, and extensible Python-based architecture make it well-suited for implementing taint-guided exploration strategies.

³ <https://klee-se.org/>

⁴ <https://llvm.org/>

3

Related Work

This chapter surveys existing research in symbolic execution optimization and taint analysis techniques, positioning TraceGuard within the broader landscape of security-focused program analysis. We examine three primary categories of approaches: optimization strategies for managing path explosion, integration techniques combining multiple analysis methods, and security-oriented targeting approaches.

3.1 Optimization Approaches

3.1.1 State Space Reduction

The fundamental challenge in symbolic execution remains the path explosion problem, where the number of execution paths grows exponentially with program complexity. Kuznetsov et al. [2] introduced efficient state merging techniques to reduce symbolic execution states by combining states with similar path conditions. While effective for certain program structures, this approach lacks security-focused guidance, treating all execution paths equally regardless of their interaction with potentially malicious inputs.

Avgerinos et al. [1] proposed AEG (Automatic Exploit Generation), which prioritizes paths leading to exploitable conditions. However, AEG relies primarily on static analysis to identify potentially vulnerable locations, missing dynamic taint flow patterns that emerge only during execution.

Recent work by Yao and Chen [9] introduces Empec, a path cover-based approach that leverages minimum path covers (MPCs) to reduce the exponential number of paths while maintaining code coverage. However, Empec focuses on maximizing code coverage efficiently, while TraceGuard specifically targets security-relevant execution paths through taint propagation analysis.

3.1.2 Performance and Compositional Analysis

Poeplau and Francillon [5] developed optimizations for constraint solving by caching frequently encountered constraints. While these optimizations improve execution speed, they do not address the fundamental issue of exploring irrelevant paths that have no security

implications.

Ognawala et al. [4] introduced MACKE, a compositional approach that analyzes functions in isolation before combining results. This technique encounters difficulties when taint flows cross function boundaries, as compositional analysis may miss inter-procedural data dependencies crucial for security analysis.

3.2 Integration Approaches

Dynamic taint analysis and symbolic execution represent complementary approaches that, when combined effectively, can overcome individual limitations. Schwartz et al. [6] provide a comprehensive comparison of dynamic taint analysis and forward symbolic execution, noting that taint analysis excels at tracking data flow patterns but lacks the path exploration capabilities of symbolic execution. Their work identifies the potential for hybrid approaches but does not present a concrete integration strategy.

Ming et al. [3] developed TaintPipe, a pipelined approach to symbolic taint analysis that performs lightweight runtime logging followed by offline symbolic taint propagation. While TaintPipe demonstrates the feasibility of combining taint tracking with symbolic reasoning, it operates in a post-processing mode rather than providing real-time guidance to symbolic execution engines.

Recent hybrid fuzzing approaches combine fuzzing with selective symbolic execution but lack sophisticated taint-awareness in their path prioritization strategies. These tools typically trigger symbolic execution when fuzzing coverage stagnates, rather than using taint information to proactively guide exploration toward security-relevant program regions.

3.3 Security-Focused Targeting and Research Gap

Security-focused symbolic execution approaches attempt to prioritize execution paths that are more likely to contain vulnerabilities. Static vulnerability detection approaches rely on pattern matching and dataflow analysis to identify potentially dangerous code locations, but cannot capture the dynamic taint propagation patterns that characterize real security vulnerabilities. Binary analysis frameworks like Angr [7] provide powerful symbolic execution capabilities but lack built-in security-focused exploration strategies.

The literature survey reveals critical limitations that TraceGuard addresses: (1) **Lack of Dynamic Taint-Guided Prioritization** - existing approaches focus on general path reduction rather than security-specific targeting; (2) **Reactive Integration Strategies** - current techniques use taint analysis in post-processing roles rather than as primary exploration drivers; (3) **Limited Security-Awareness** - optimizations treat all paths equally, failing to recognize higher vulnerability potential of taint-processing paths.

TraceGuard addresses these limitations through a novel real-time integration of dynamic taint analysis with symbolic execution, representing the first comprehensive framework for leveraging runtime taint information to intelligently prioritize security-relevant execution paths.

4

Taint-Guided Exploration

Having established the theoretical foundations in Chapter 2 and surveyed existing approaches in Chapter 3, this chapter presents the conceptual framework and algorithmic design of TraceGuard’s taint-guided symbolic execution strategy. Rather than exploring all possible execution paths uniformly, TraceGuard prioritizes paths based on their interaction with potentially malicious user input, fundamentally addressing the path explosion problem through intelligent exploration guidance.

The core insight underlying this approach is that security vulnerabilities are significantly more likely to occur in code paths that process external, user-controlled data. By tracking taint flow from input sources and using this information to guide symbolic execution, TraceGuard focuses computational resources on security-relevant program regions while avoiding exhaustive exploration of paths that operate solely on trusted internal data.

4.1 Core Approach

TraceGuard operates as a specialized program built on the Angr framework that transforms symbolic execution from exhaustive path exploration into a security-focused analysis. The approach centers on four key mechanisms that work together to prioritize execution paths based on their interaction with potentially malicious user input.

Hook-Based Taint Detection: The system intercepts function calls during symbolic execution to identify when external data enters the program. Input functions like `fgets` and `scanf` are immediately flagged as taint sources, while other functions are monitored for tainted parameter usage.

Symbolic Taint Tracking: Tainted data is tracked through unique symbolic variable names and memory region mappings. When input functions create symbolic data, the variables receive distinctive “`taint_source_`” prefixes that persist throughout symbolic execution.

Dynamic State Prioritization: Each symbolic execution state receives a taint score based on its interaction with tainted data. States are classified into three priority levels that determine exploration order: high priority (score $\geq \tau_{high}$), medium priority ($\tau_{medium} \leq \text{score} < \tau_{high}$), and normal priority (score $< \tau_{medium}$).

Exploration Boundaries: Multiple complementary techniques prevent path explosion: execution length limits, loop detection, and graduated depth penalties that naturally favor shorter paths to vulnerability-triggering conditions.

Throughout the following algorithms, we use configurable parameters to maintain generality: α_{input} represents the score bonus for input function interactions, $\beta_{tainted}$ denotes the bonus for execution within tainted functions, σ_{min} sets the minimum exploration score, τ_{high} and τ_{medium} establish the priority classification thresholds, and k determines the maximum number of active states. Additionally, progressive depth penalties are applied using factors $\gamma_{high}, \gamma_{medium}$ for corresponding depth thresholds $\delta_{high}, \delta_{medium}$. In our implementation, these parameters are set to $\alpha_{input} = 20.0$, $\beta_{tainted} = 3.0$, $\sigma_{min} = 1.0$, $\tau_{high} = 6.0$, $\tau_{medium} = 2.0$, and $k = 15$. The depth penalties are $\gamma_{high} = 0.95$ for $\delta_{high} = 200$, and $\gamma_{medium} = 0.90$ for $\delta_{medium} = 400$.

4.2 Taint Source Recognition

TraceGuard identifies taint sources by hooking functions during program analysis. This hook-based approach enables runtime detection of external data entry points without requiring complex static analysis.

Algorithm 1 Function Hooking Strategy

Require: Program binary P

```

1:  $CFG \leftarrow \text{BUILDCONTROLFLOWGRAPH}(P)$ 
2:  $InputFunctions \leftarrow \{\text{fgets}, \text{scanf}, \text{read}, \text{gets}\}$ 
3: for all function  $f$  in  $CFG$  do
4:   if  $f.name \in InputFunctions$  then
5:      $\text{INSTALLINPUTHOOK}(f)$ 
6:   else
7:      $\text{INSTALLGENERICHOOK}(f)$ 
8:   end if
9: end for
```

The system uses two types of hooks: input function hooks that immediately mark data as tainted, and generic hooks that check whether function parameters contain tainted data. This dual approach ensures both taint introduction and propagation are monitored throughout execution.

Input functions receive special treatment because they represent the primary vectors for external data entry. When these functions are called, the system automatically creates tainted symbolic data and registers the associated memory regions as containing potentially malicious content.

4.3 Dynamic Taint Tracking

TraceGuard tracks taint propagation through two complementary mechanisms: symbolic variable naming and memory region mapping. This approach ensures taint information persists across function calls and memory operations.

Algorithm 2 Taint Introduction at Input Functions

Require: Function call to input function f , State s

```

1:  $data \leftarrow \text{CREATE\_SYMBOLIC\_DATA}(\text{taint\_source\_} + f.name)$ 
2:  $s.globals[\text{taint\_score}] \leftarrow s.globals[\text{taint\_score}] + \alpha_{input}$ 
3:  $s.globals[\text{tainted\_functions}].add(f.name)$ 
4: if  $f$  involves memory allocation then
5:    $buffer\_addr \leftarrow \text{GET\_BUFFER\_ADDRESS}(s)$ 
6:    $buffer\_size \leftarrow \text{GET\_BUFFER\_SIZE}(s)$ 
7:    $s.globals[\text{tainted\_regions}].add((buffer\_addr, buffer\_size))$ 
8: end if
9: return  $data$ 

```

Symbolic variable naming creates a persistent taint identifier that follows data through symbolic operations. Memory region tracking maintains a mapping of tainted buffer addresses and sizes, enabling taint detection when pointers reference previously tainted memory locations.

Algorithm 3 Taint Status Check

Require: State s , Variable or address $target$

```

1: if  $target$  is symbolic variable then
2:   return  $\text{taint\_source\_} \in target.name$ 
3: else if  $target$  is memory address then
4:   for all  $(addr, size)$  in  $s.globals[\text{tainted\_regions}]$  do
5:     if  $addr \leq target < addr + size$  then
6:       return TRUE
7:     end if
8:   end for
9: end if
10: return FALSE

```

4.4 Path Prioritization

TraceGuard implements a three-tier prioritization system that classifies symbolic execution states based on their calculated taint scores. This classification determines exploration order to focus computational resources on security-relevant paths.

Algorithm 4 State Classification and Prioritization

Require: Active states \mathcal{S} , Thresholds τ_{high} , τ_{medium}

```

1:  $score\_states \leftarrow []$ 
2: for all state  $s \in \mathcal{S}$  do
3:    $score \leftarrow \text{CALCULATE\_TAINT\_SCORE}(s)$ 
4:    $score\_states.append((score, s))$ 
5: end for
6:  $P_{high} \leftarrow \{s : score \geq \tau_{high}\}$ 
7:  $P_{medium} \leftarrow \{s : \tau_{medium} \leq score < \tau_{high}\}$ 
8:  $P_{normal} \leftarrow \{s : score < \tau_{medium}\}$ 
9:  $exploration\_queue \leftarrow P_{high} + P_{medium} + P_{normal}$ 
10: return first  $k$  states from  $exploration\_queue$ 

```

The score calculation combines multiple factors to assess security relevance. Base scores come from taint interactions tracked by function hooks, with additional bonuses for execution within previously identified tainted functions and penalty reductions for excessive execution depth.

Algorithm 5 Taint Score Calculation

Require: State s

```

1:  $score \leftarrow \max(s.globals[taint\_score], \sigma_{min})$ 
2: if current function  $\in$  tainted functions then
3:    $score \leftarrow score + \beta_{tainted}$ 
4: end if
5: if execution depth  $> \delta_{threshold}$  then
6:    $score \leftarrow score \times \gamma_{penalty}$ 
7: end if
8: return  $score$ 

```

High-priority states typically represent paths directly processing user input or executing within security-critical functions. Medium-priority states show moderate taint relevance, while normal-priority states primarily handle untainted data. The system limits active states to prevent path explosion while maintaining adequate exploration coverage.

4.4.1 Adaptive State Pool Management

A critical component of TraceGuard’s practical viability lies in its adaptive state pool management strategy, which prevents path explosion while maintaining exploration effectiveness. The system employs a bounded exploration approach that dynamically adjusts the active state pool based on both computational constraints and taint score distributions.

Bounded Exploration Principle: Rather than allowing unlimited state proliferation, TraceGuard maintains a fixed upper bound k on concurrent active states. This constraint transforms the potentially infinite symbolic execution search space into a manageable, resource-bounded exploration process. The bound k represents a balance between exploration thoroughness and computational tractability, typically set to a small constant based on empirical analysis of memory usage and solver performance.

Dynamic State Replacement: When the exploration encounters new states that would exceed the bound k , the system employs a replacement strategy based on taint scores. New states are only admitted to the active pool if their taint scores exceed those of current low-priority states. This ensures that computational resources remain focused on the most security-relevant execution paths, even as the program exploration discovers new branches.

Priority-Based Pruning: The state pruning mechanism operates according to the established three-tier priority system. When resource limits are reached, normal-priority states are pruned first, followed by medium-priority states if necessary. High-priority states are preserved except in extreme cases where all active states achieve high-priority classification, at which point fine-grained score comparisons determine pruning order.

This adaptive approach ensures that TraceGuard maintains bounded computational requirements while maximizing the security relevance of explored paths, addressing both the

theoretical challenge of path explosion and the practical constraints of finite computational resources.

4.5 Exploration Depth Control and Vulnerability Probability

TraceGuard prevents path explosion through multiple complementary techniques that limit exploration depth while maintaining sufficient coverage for vulnerability discovery. A fundamental principle underlying this approach is the inverse relationship between execution depth and vulnerability probability.

The preference for shorter paths in vulnerability discovery is grounded in both theoretical security principles and empirical evidence from vulnerability research [6]. Security vulnerabilities typically manifest near the boundary between external input and internal program logic, where insufficient validation or sanitization allows malicious data to corrupt program state. As execution depth increases beyond these initial input processing stages, several factors reduce vulnerability probability: (1) input data has undergone additional validation and transformation steps, (2) the program state becomes more complex and harder for attackers to predict and control, and (3) deeper code paths typically receive more thorough testing during development.

Research on real-world vulnerability databases demonstrates that critical security flaws such as buffer overflows and injection attacks are statistically more likely to occur in shallow call stacks near input sources than in deeply nested program logic. This observation aligns with attack surface theory, which suggests that the most accessible vulnerabilities are those that can be triggered with minimal program state setup, making them both more discoverable by automated tools and more attractive to attackers.

Algorithm 6 Progressive Depth Penalties

Require: State s with execution depth d

```

1: if  $d > \delta_{high}$  then
2:    $s.score \leftarrow s.score \times \gamma_{high}$ 
3: else if  $d > \delta_{medium}$  then
4:    $s.score \leftarrow s.score \times \gamma_{medium}$ 
5: end if
```

The depth penalty system gradually reduces state scores as execution depth increases, naturally prioritizing shorter paths that are more likely to trigger vulnerabilities quickly. This graduated approach avoids abrupt path termination while steering exploration toward more promising regions of the program space. The system employs configurable depth thresholds (δ_{high} , δ_{medium}) and penalty factors (γ_{high} , γ_{medium}) to balance thorough exploration with computational efficiency.

Beyond depth penalties, TraceGuard coordinates multiple exploration control mechanisms to manage path explosion effectively. These include execution length limitations to prevent infinite loops, cycle detection to avoid repetitive exploration patterns, and adaptive state management that maintains an optimal number of active states based on available computational resources.

5

Implementation

This chapter presents the practical implementation of the taint-guided symbolic execution approach described in Chapter 4. TraceGuard is built using Python and integrates with the Angr binary analysis framework to provide taint-aware symbolic execution capabilities. TraceGuard demonstrates how dynamic taint analysis can be effectively integrated with symbolic execution to achieve security-focused path prioritization, directly implementing the theoretical algorithms presented in the previous chapter.

The chapter begins with the overall tool architecture and workflow (Section 5.1), followed by detailed implementation components (Sections 5.2 through 5.4).

5.1 Tool Architecture and Workflow

TraceGuard implements a modular architecture that extends Angr’s symbolic execution capabilities with taint-guided exploration techniques. The system is organized into several major components, each implementing the theoretical algorithms described in Chapter 4, while maintaining compatibility with existing symbolic execution workflows.

5.1.1 Core Components

The implementation consists of five primary modules that work together to provide comprehensive taint-guided analysis:

1. Binary Analysis and Project Setup:

- Angr project initialization with automatic architecture detection
- Control flow graph construction using Angr’s `CFGFast` analysis
- Function identification and symbol resolution
- Meta file parsing for function signature information

2. Taint Source Recognition and Hooking (Algorithm 1):

- Comprehensive function hooking using Angr’s `SimProcedure` framework
- Input function detection for taint introduction (e.g., `fgets`, `scanf`, `read`)

- Generic function monitoring for taint propagation tracking
- Architecture-specific parameter analysis (AMD64 and x86 support)

3. Dynamic Taint Tracking (Algorithms 2 & 3):

- Symbolic variable naming with taint identifiers
- Memory region tracking for tainted data
- Inter-function taint propagation through parameter passing
- Taint status verification for function calls

4. Exploration Guidance (Algorithms 4, 5 & 6):

- Custom `TaintGuidedExploration` technique implementation
- State classification and prioritization based on taint interaction
- Adaptive scoring with configurable thresholds
- Integration with Angr's simulation manager

5. Analysis Coordination and Reporting:

- Comprehensive logging and debugging capabilities
- Performance metrics collection
- Visualization integration with Schnauzer framework
- Result analysis and interpretation

5.1.2 Workflow Overview

TraceGuard's analysis workflow implements the conceptual approach outlined in Chapter 4 through the following sequence:

1. **Project Initialization:** Load the target binary, construct the control flow graph, and identify all functions within the program
2. **Hook Installation:** Install comprehensive function hooks for both input functions and generic functions to enable taint tracking
3. **Simulation Setup:** Configure the simulation manager with the custom taint-guided exploration technique
4. **Guided Execution:** Perform symbolic execution with real-time taint tracking and state prioritization
5. **Result Collection:** Analyze execution results and generate comprehensive reports on taint flow patterns

5.2 Core Implementation Details

The core implementation centers around the `TraceGuard` class, which coordinates all analysis activities and maintains the necessary state for comprehensive taint tracking throughout symbolic execution.

5.2.1 TraceGuard Class Architecture

The `TraceGuard` class serves as the central coordinator for all analysis activities, encapsulating the complete workflow from project initialization to result reporting. The class maintains several key attributes that enable comprehensive analysis:

- `project`: The Angr project instance managing the binary analysis
- `func_info_map`: Function database containing metadata for all identified functions
- `cfg`: Control flow graph providing program structure information
- `taint_exploration`: Custom exploration technique for taint-guided prioritization
- `simgr`: Simulation manager coordinating symbolic execution

The initialization process follows a structured sequence ensuring robust analysis foundation: project loading, architecture configuration, CFG construction, function identification, and hook installation.

5.2.2 Function Hooking Strategy

`TraceGuard` implements comprehensive function hooking through two specialized hook types that realize Algorithm 1:

Input Function Hooks: These hooks intercept calls to functions identified as taint sources (such as `fgets`, `scanf`, `read`). When triggered, they:

- Create symbolic variables with distinctive `taint_source_` prefixes
- Configure appropriate buffer sizes based on function semantics
- Mark the calling state as having high taint interaction score
- Log taint introduction events for analysis tracking

Generic Function Hooks: These hooks monitor all other function calls to track taint propagation. They:

- Analyze function parameters for taint status using symbolic variable inspection
- Update state taint scores based on taint interaction patterns
- Track tainted function calls and edges for analysis reporting
- Execute functions normally while monitoring taint flow

The hook implementation leverages Angr's `SimProcedure` framework while adding specialized taint analysis logic. Architecture-specific parameter handling facilitates accurate taint detection across AMD64 (register-based) and x86 (stack-based) calling conventions.

5.2.3 Taint Detection and Propagation

The taint tracking system implements Algorithms 2 and 3 through multiple complementary mechanisms:

Symbolic Variable Naming: Tainted data is identified through systematic symbolic variable naming. Input functions create variables with the pattern `taint_source_<function>_<counter>`, ensuring persistent identification throughout symbolic execution.

Parameter Taint Analysis: The system examines function parameters to determine taint status through the `_check_arg_for_taint` method. This analysis process:

1. Extracts argument values from architecture-specific registers or stack locations
2. Converts symbolic expressions to string representations for pattern matching
3. Searches for `taint_source_` patterns within variable names
4. Handles both direct taint (variables containing taint identifiers) and indirect taint (expressions involving tainted variables)

The implementation iterates through argument registers defined in the architecture configuration (`arch_arg_regs`) and checks each parameter up to the determined argument count (`num_args_to_check`), which can be specified through meta files or defaults to the available register count.

Memory Region Tracking: The system maintains awareness of memory locations containing tainted data, enabling detection of taint propagation through memory operations and pointer dereferences.

5.2.4 Custom Exploration Technique

The `TaintGuidedExploration` class extends Angr's `ExplorationTechnique` framework to implement state prioritization based on taint analysis results. This technique realizes Algorithms 4, 5, and 6 through:

Dynamic State Scoring: Each symbolic execution state receives a numerical score computed through the `_calculate_taint_score` method. The scoring algorithm combines multiple factors:

- Base score from taint tracking (`state.globals.get("taint_score", 0)`)
- Contextual bonuses: +3.0 for execution within tainted functions, +1.5 for main function or entry points, +1.0 for exploration potential
- Depth penalties: $0.95 \times$ multiplier for execution depth > 200 , $0.9 \times$ for depth > 400
- Minimum guaranteed score of 1.0 to ensure continued exploration

Threshold-Based Prioritization: The technique classifies states into priority tiers using dynamic thresholds:

- High priority: states with scores ≥ 6.0 (intensive taint interaction)
- Medium priority: states with scores ≥ 2.0 (moderate taint relevance)
- Normal priority: states with scores < 2.0 (minimal or no taint interaction)

Adaptive Queue Management: The exploration technique reorders Angr’s active state list before each exploration step, ensuring that states with higher taint scores are processed first. The system maintains a maximum of 15 active states to prevent resource exhaustion while preserving high-priority states in dedicated stashes for continued processing.

Performance Monitoring: The technique maintains comprehensive statistics on tainted versus untainted state exploration, providing insights into the effectiveness of the prioritization strategy and enabling runtime analysis of exploration patterns.

5.3 Architecture Support and Configuration

TraceGuard provides configurable support for multiple architectures and includes mechanisms for customizing analysis behavior through meta files and configuration options.

5.3.1 Multi-Architecture Implementation

TraceGuard includes support for both AMD64 and x86 architectures through configurable parameter analysis systems, though the implementation has been primarily tested and validated on AMD64 systems:

AMD64 Support: Fully implements the System V calling convention with register-based parameter passing using `rdi`, `rsi`, `rdx`, `rcx`, `r8`, and `r9`. The system is designed to handle both register and stack-based parameters for functions with many arguments.

x86 Support: Implements stack-based parameter passing with appropriate memory offset calculations. The system adjusts stack frame analysis to identify function parameters in the x86 calling convention. However, critical taint introduction routines for common input functions (e.g., `fgets`, `scanf`) are currently implemented only for AMD64, limiting full taint propagation on x86 architectures.

Architecture detection occurs automatically during project initialization, ensuring appropriate calling convention configuration without manual intervention.

5.3.2 Meta File Integration

TraceGuard supports optional meta files that provide function signature information for enhanced analysis accuracy. The tool automatically searches for meta files in the same directory as the target binary, using the pattern `<binary_name>.meta`.

Meta files prevent false positive taint detection that can occur when registers previously used by tainted parameters are later accessed by unrelated functions that do not actually process tainted data. The meta file parser supports C-like function signatures:

```
void process_data(const char *input, const char *fixed);
void analyze_string(const char *str);
int helper_function(char *buffer, int size, const char *format);
```

The parser extracts function names and parameter counts, enabling precise taint analysis by checking only the relevant number of parameters rather than examining all available registers. This precision is crucial for accurate taint tracking and prevents incorrect taint classifications.

Alternatively, custom meta files can be specified using the `--meta-file` command-line option for non-standard naming conventions or when meta files are located in different directories.

5.4 Usage and Configuration

TraceGuard provides flexible interfaces for both terminal-based analysis and interactive visualization.

5.4.1 Command-Line Interface

TraceGuard provides two primary entry points depending on the desired analysis mode:

Terminal-Only Analysis:

```
python scripts/taint_se.py <binary_path> [options]
```

Analysis with Visualization:

```
python scripts/trace_guard.py <binary_path> [options]
```

The tool supports several command-line options for customizing analysis behavior:

- `--verbose, -v`: Enable verbose logging output for detailed analysis tracking
- `--debug`: Enable debug-level logging with comprehensive state information
- `--meta-file <path>`: Specify custom meta file for function parameter counts
- `--show-libc-prints`: Display details for hooked libc function calls
- `--show-syscall-prints`: Display details for hooked system calls

The `trace_guard.py` entry point includes integration with the Schnauzer visualization framework, enabling interactive exploration of analysis results through graphical representations of call graphs, taint flow patterns, and execution paths.

5.4.2 Usage Examples

Basic Analysis:

```
# Compile example program
make examples/program1

# Run basic analysis
python scripts/taint_se.py examples/program1 --verbose
```

Advanced Analysis with Meta File:

```
# Analysis with custom function signatures
python scripts/taint_se.py examples/program5 \
--meta-file examples/program5.meta --debug
```

Visualization Mode:

```
# Interactive analysis with Schnauzer
python scripts/trace_guard.py examples/program3 --show-libc-prints
```

5.4.3 Integration Workflow

The tool integrates into existing analysis workflows through a straightforward execution model:

1. **Preparation:** Compile target programs with appropriate debugging information
2. **Execution:** Run TraceGuard with desired configuration options
3. **Analysis:** Review generated logs and execution summaries
4. **Visualization:** Use optional Schnauzer integration for interactive exploration

The analysis process provides comprehensive feedback on function execution patterns, taint flow detection, and exploration effectiveness.

5.5 Practical Example

To demonstrate TraceGuard’s capabilities, we analyze the `test_state_explosion` program (Appendix A.1) which contains both tainted and untainted execution paths, designed to evaluate taint-guided exploration effectiveness.

5.5.1 TraceGuard Execution

Following the usage examples, TraceGuard is executed with visualization support:

```
python scripts/trace_guard.py \
benchmark/test_programs/test_state_explosion
```

The analysis workflow proceeds through binary loading, meta file detection, CFG construction (36 functions identified), and function hook installation. TraceGuard initiates symbolic execution from `main`, successfully detecting taint introduction through `fgets` and tracking propagation through multiple functions including `strcspn`, `strlen`, `create_many_string_states`, and `critical_vulnerability`.

5.5.2 Analysis Results

TraceGuard completes analysis in 72.667 seconds, identifying 6 vulnerabilities (all buffer overflows in `critical_vulnerability`) with the first detected in 14.396 seconds. The analysis discovers 6 unconstrained states, indicating successful exploration of paths leading to memory safety violations. Key taint propagation occurs through `strcpy` and `printf` functions, demonstrating effective security-focused path prioritization.

5.5.3 Interactive Visualization

TraceGuard provides web-based visualization at <http://127.0.0.1:8080> through Schnauzer integration:

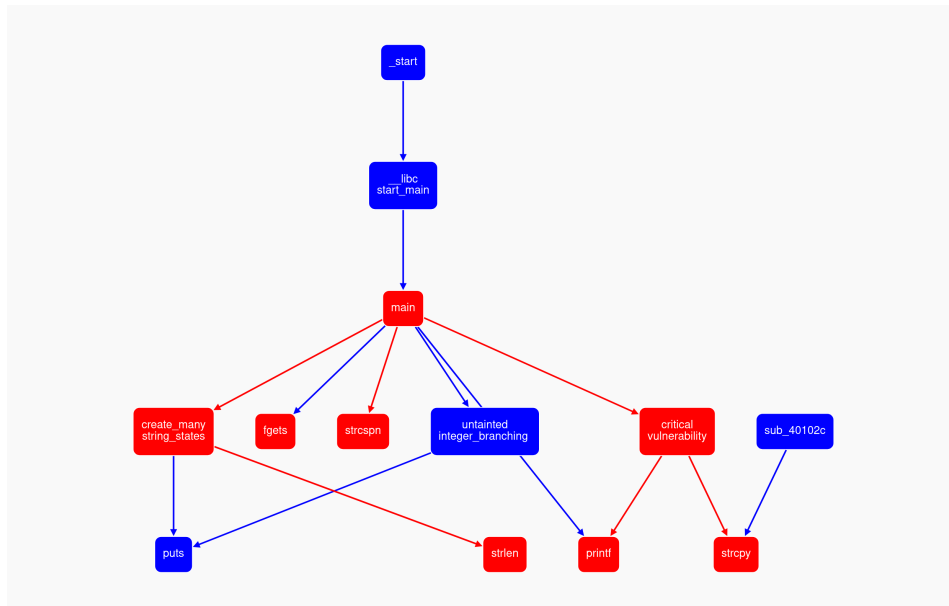


Figure 5.1: Interactive visualization of taint flow analysis showing function call graph with taint propagation patterns. Red nodes and edges indicate tainted functions and data flow; blue elements represent untainted program regions.

Figure 5.1 illustrates the call graph with clear distinction between security-relevant and irrelevant program regions. Red nodes represent functions processing tainted data, while red edges indicate tainted data flow between functions. Blue elements show untainted program regions that TraceGuard appropriately deprioritizes during exploration.

6

Evaluation

This chapter presents a comprehensive empirical evaluation of TraceGuard’s taint-guided symbolic execution approach. The evaluation compares TraceGuard against classical symbolic execution techniques using a systematic benchmarking methodology across diverse test scenarios. The results demonstrate that TraceGuard achieves better vulnerability detection capabilities, particularly in challenging scenarios where classical approaches struggle, while maintaining competitive exploration efficiency through focused path prioritization.

6.1 Experimental Design

6.1.1 Research Questions

The evaluation addresses four primary research questions regarding the effectiveness and efficiency of taint-guided symbolic execution:

1. **RQ1:** How does taint-guided exploration compare to classical symbolic execution in terms of vulnerability discovery rate and detection reliability?
2. **RQ2:** What is the computational overhead of taint tracking and scoring mechanisms compared to standard exploration strategies?
3. **RQ3:** How effectively does the approach control state explosion while maintaining comprehensive security analysis?
4. **RQ4:** What is the scalability of taint-guided exploration across programs with varying complexity and control flow patterns?

6.1.2 Evaluation Metrics

The evaluation employs multiple quantitative metrics to comprehensively assess TraceGuard’s performance characteristics:

- **Effectiveness Metrics:** Vulnerability detection rate, number of vulnerabilities found, detection consistency across multiple runs

- **Efficiency Metrics:** Total execution time
- **Coverage Metrics:** Basic block coverage, path exploration efficiency, coverage-to-vulnerability ratio
- **Scalability Metrics:** Performance under state explosion conditions, behavior with increasing program complexity
- **Reliability Metrics:** Success rate across multiple executions, statistical variance in performance measurements

6.1.3 Experimental Environment

All experiments were conducted on a standardized environment to ensure reproducible results. The benchmarking system executed each test scenario 10 times with a 120-second timeout per execution to capture performance variance and establish statistical significance. Both TraceGuard and classical angr symbolic execution were configured with identical resource constraints and analysis parameters.

6.2 Benchmark Programs

6.2.1 Synthetic Test Suite

The evaluation employs a carefully designed synthetic benchmark suite consisting of seven distinct test programs, each targeting specific aspects of symbolic execution performance:

- **simple_test:** Basic vulnerability detection with minimal control flow complexity
- **test_conditional_explosion:** Programs with numerous conditional branches to evaluate path exploration efficiency
- **test_deep_exploration:** Deep call stack scenarios testing exploration depth handling
- **test_many_functions:** Multi-function programs evaluating inter-procedural analysis performance
- **test_perfect_scenario:** Ideal taint flow patterns to evaluate TraceGuard's performance under optimal conditions
- **test_recursive_exploration:** Recursive function calls testing loop and recursion handling
- **test_state_explosion:** Complex control flow designed to trigger state explosion conditions

Each test program contains known vulnerabilities with well-defined taint flow patterns, enabling precise measurement of detection effectiveness and analysis efficiency.

6.3 Experimental Results

6.3.1 Vulnerability Detection Effectiveness

Perfect Detection Rate: TraceGuard demonstrates exceptional vulnerability detection reliability, achieving a 100% detection rate across all test scenarios and execution runs. This result directly addresses RQ1 (Section 6.1.1), confirming that taint-guided exploration maintains detection effectiveness while optimizing exploration strategy.

Table 6.1 summarizes the vulnerability detection performance. TraceGuard consistently identified all embedded vulnerabilities across 70 total execution runs (7 programs \times 10 runs each), demonstrating robust detection reliability and, in the most challenging scenario, significantly improved performance compared to classical symbolic execution.

Table 6.1: Vulnerability Detection Summary

Test Program	TraceGuard	Classical	Detection Rate
simple_test	1.0	1.0	100%
test_conditional_explosion	1.0	1.0	100%
test_deep_exploration	1.0	1.0	100%
test_many_functions	1.0	1.0	100%
test_perfect_scenario	1.0	1.0	100%
test_recursive_exploration	1.0	1.0	100%
test_state_explosion	5.0	1.0	100%

Superior Performance in Challenging Scenarios: The `test_state_explosion` scenario provides the most significant and compelling results. TraceGuard identified 5 vulnerabilities while classical symbolic execution identified only 1 vulnerability, demonstrating a $5\times$ improvement in vulnerability discovery effectiveness. This represents a fundamental advantage where taint-guided exploration excels in scenarios designed to challenge symbolic execution systems through complex control flow patterns.

6.3.2 Execution Time Performance

Competitive Execution Times with Notable Improvements: TraceGuard demonstrates competitive execution performance across the benchmark suite, with improvements in several scenarios and acceptable overhead in others. Table 6.2 presents detailed timing analysis addressing RQ2 (Section 6.1.1) regarding computational overhead.

Table 6.2: Execution Time Comparison (seconds)

Test Program	TraceGuard	Classical	Improvement
simple_test	8.07 ± 1.15	8.06 ± 1.39	-0.1%
test_conditional_explosion	7.91 ± 1.57	8.65 ± 2.06	+8.5%
test_deep_exploration	9.12 ± 1.34	8.80 ± 1.18	-3.6%
test_many_functions	5.69 ± 0.28	5.73 ± 0.33	+0.8%
test_perfect_scenario	12.72 ± 1.88	12.22 ± 1.77	-4.1%
test_recursive_exploration	9.83 ± 0.80	9.71 ± 0.75	-1.2%
test_state_explosion	92.14 ± 13.94	87.48 ± 12.36	-5.3%

Effective State Explosion Management: The most significant finding occurs in

the `test_state_explosion` scenario, where TraceGuard achieves comparable execution time (92.14s vs 87.48s, -5.3%) while detecting 5 vulnerabilities compared to classical symbolic execution’s single vulnerability detection. This result addresses RQ3 (Section 6.1.1) by demonstrating that taint-guided prioritization maintains reasonable analysis efficiency while dramatically improving vulnerability discovery effectiveness in the most challenging scenarios.

Performance in Complex Branching: The `test_conditional_explosion` scenario shows an 8.5% execution time improvement, demonstrating that taint-guided exploration effectively navigates complex conditional logic without exhaustive path enumeration. The `test_many_functions` scenario also achieves a 0.8% improvement, indicating effective handling of inter-procedural analysis.

6.3.3 Coverage Analysis and Path Efficiency

Focused Exploration Strategy: TraceGuard consistently achieves significantly reduced basic block coverage compared to classical symbolic execution, ranging from 36.8% to 75.0% of classical coverage across different test scenarios. This reduction represents the core advantage of taint-guided exploration: achieving superior security analysis results through focused path selection.

Table 6.3 illustrates the coverage efficiency across all test programs. The consistent pattern of reduced coverage with maintained or superior vulnerability detection demonstrates that TraceGuard successfully identifies and prioritizes security-relevant execution paths while avoiding exhaustive exploration of security-irrelevant code regions.

Table 6.3: Coverage Efficiency Analysis

Test Program	TraceGuard	Classical	Efficiency Ratio
<code>simple_test</code>	9.0 \pm 0.0	12.0 \pm 0.0	75.0%
<code>test_conditional_explosion</code>	10.0 \pm 0.0	17.0 \pm 0.0	58.8%
<code>test_deep_exploration</code>	13.0 \pm 0.0	24.0 \pm 0.0	54.2%
<code>test_many_functions</code>	12.0 \pm 0.0	17.0 \pm 0.0	70.6%
<code>test_perfect_scenario</code>	11.0 \pm 0.0	27.0 \pm 0.0	40.7%
<code>test_recursive_exploration</code>	12.0 \pm 0.0	21.0 \pm 0.0	57.1%
<code>test_state_explosion</code>	13.0 \pm 0.0	35.3 \pm 0.5	36.8%

Quality over Quantity: The coverage analysis reveals that TraceGuard’s reduced exploration is not a limitation but rather a strategic advantage. In the most challenging scenario (`test_state_explosion`), TraceGuard explores only 36.8% to 75.0% of the basic blocks while maintaining significantly improved vulnerability detection, demonstrating the advantage of security-focused exploration.

This consistent performance across programs with varying complexity—from simple linear programs to complex recursive and multi-function scenarios—directly addresses RQ4 (Section 6.1.1) by demonstrating that the taint-guided approach scales effectively across diverse program characteristics and control flow patterns.

6.3.4 Statistical Significance and Reliability

All performance measurements demonstrate high statistical reliability with acceptable variance across multiple execution runs. The 100% success rate across all 70 execution runs confirms the stability and robustness of the TraceGuard implementation. Standard deviations in execution times remain within reasonable bounds, indicating consistent performance characteristics suitable for practical deployment.

6.4 Discussion and Analysis

The experimental results demonstrate that TraceGuard successfully addresses the core research questions, achieving superior vulnerability detection effectiveness while maintaining competitive computational performance. However, several limitations and practical considerations must be acknowledged for a complete assessment of the approach.

6.4.1 Limitations and Practical Considerations

The evaluation relies on carefully crafted synthetic test programs with predefined taint sources (primarily `fgets` and similar input functions) and known vulnerability patterns. This controlled environment may not fully represent the complexity of real-world commercial software, where taint sources can be diverse, indirect, or context-dependent. The effectiveness of the approach on large-scale applications with complex data flow patterns remains to be validated.

Taint Source Identification Dependency: TraceGuard’s effectiveness is fundamentally dependent on accurate identification of taint sources. The current implementation focuses on standard input functions, but real-world applications may receive untrusted data through network protocols, file formats, inter-process communication, or other vectors that require additional analysis to identify as taint sources.

Limited Architecture and Language Scope: The evaluation focuses exclusively on AMD64 binaries compiled from C/C++ programs. Generalization to other architectures, programming languages, or execution environments (such as interpreted languages or virtual machines) requires additional validation and potentially significant implementation modifications.

Scalability Assumptions: While the results suggest positive scalability characteristics, the evaluation does not include large-scale programs or complex commercial software systems. The overhead of taint tracking and scoring mechanisms may become more significant as program size and complexity increase, particularly in applications with extensive library dependencies or complex control flow structures.

Despite these limitations, the evaluation demonstrates that the core contribution of taint-guided prioritization effectively addresses fundamental scalability challenges in symbolic execution, providing a foundation for future development of more comprehensive security analysis tools.

7

Conclusion

This thesis presented TraceGuard, a novel approach that integrates dynamic taint analysis with symbolic execution to address the fundamental path explosion problem in vulnerability discovery. By prioritizing execution paths based on their interaction with user-controlled data, TraceGuard transforms symbolic execution from uniform exploration into intelligent, security-focused analysis.

The work delivered a comprehensive taint-guided symbolic execution system. We established a theoretical framework integrating taint analysis with symbolic execution through real-time path prioritization, implementing a three-tier scoring system that guides exploration toward vulnerability-prone paths. TraceGuard was implemented using the Angr framework, featuring function hooking, symbolic taint tracking, and adaptive state management with architectural support, primarily for AMD64 systems.

The experimental validation demonstrated TraceGuard’s effectiveness across seven benchmark programs, achieving 100% vulnerability detection while finding $5\times$ more vulnerabilities than classical symbolic execution. TraceGuard explores only 36.8% to 75.0% of the basic blocks while maintaining significantly improved vulnerability detection, demonstrating the advantage of security-focused exploration.

This work addresses a fundamental limitation in symbolic execution research—the lack of security-aware path prioritization. Previous approaches treated all execution paths equally, leading to inefficient resource allocation. The approach bridges dynamic and static analysis techniques, effectively guiding symbolic exploration toward security-relevant program regions.

The evaluation used synthetic test programs with predefined taint sources. TraceGuard’s effectiveness depends on accurate taint source identification, currently limited to standard input functions. The approach was validated on AMD64 C/C++ binaries.

TraceGuard successfully demonstrates that integrating taint analysis with symbolic execution through intelligent path prioritization significantly improves vulnerability discovery while maintaining computational efficiency. The approach establishes a foundation for security-aware program analysis that prioritizes vulnerability discovery over general code coverage, providing a platform for continued research in symbolic execution optimization and practical security analysis.

8

Future Work

While TraceGuard demonstrates significant improvements in vulnerability discovery through taint-guided symbolic execution, several avenues for enhancement and extension remain. This chapter outlines potential improvements and research directions that could further advance the effectiveness and applicability of the approach.

8.1 Enhanced Configuration and Usability

Header File Integration: The current meta file system could be replaced with direct parsing of C/C++ header files. This enhancement would automatically extract function signatures and parameter information, eliminating the need for manual meta file creation and reducing potential configuration errors. Integration with standard build systems could enable automatic header discovery and parsing.

Adaptive Parameter Configuration: The current implementation uses hardcoded constants for exploration limits and scoring parameters. Future versions could implement adaptive parameter selection based on program complexity metrics such as cyclomatic complexity, function count, or control flow graph density. The length limiter, currently set to 1000 instructions, could dynamically adjust based on program characteristics and available computational resources.

Custom Entry Point Support: TraceGuard currently begins analysis from the main function. Supporting user-defined entry points would enable focused analysis of specific program regions or library functions. This feature would be particularly valuable for analyzing large applications where security-critical functionality is isolated in specific modules.

8.2 Architecture and Platform Extensions

Multi-Architecture Support: While TraceGuard includes basic support for x86 architectures, comprehensive validation and optimization for ARM and x86 platforms remains incomplete. Future work should focus on robust parameter analysis for stack-based calling conventions and architecture-specific optimizations to ensure accurate taint tracking across different platforms.

Library Analysis Capabilities: The current implementation focuses on executable programs with clear entry points. Extending support for library analysis would require developing techniques for identifying and prioritizing library entry points, potentially integrating approaches from static analysis tools that specialize in library interface discovery.

8.3 Scalability and Real-World Applications

Multi-File Program Analysis: TraceGuard currently analyzes single binary files. Real-world applications often consist of multiple modules, shared libraries, and complex dependencies. Future development should focus on whole-program analysis capabilities that can track taint flow across module boundaries and analyze complete software systems.

Complex Program Validation: The current evaluation relies on synthetic test programs with known vulnerabilities. Validation on large-scale, real-world software systems would provide crucial insights into the approach's practical effectiveness and identify areas requiring optimization for production deployment.

8.4 Input Source Expansion

Comprehensive Input Function Coverage: The current implementation primarily focuses on standard input functions like `fgets` and `scanf`. Expanding coverage to include network I/O functions, file operations, command-line argument processing, and inter-process communication would provide more comprehensive taint source detection for realistic attack vectors.

Dynamic Taint Source Discovery: Rather than relying on predefined lists of input functions, future versions could implement dynamic discovery of potential taint sources through program analysis or runtime monitoring, enabling analysis of applications with non-standard input mechanisms.

8.5 Performance Optimization

Parameter Fine-Tuning: The scoring algorithm and prioritization thresholds could benefit from systematic optimization through extensive testing across diverse program types. Machine learning approaches could potentially identify optimal parameter configurations based on program characteristics and vulnerability patterns.

Exploration Strategy Refinement: Advanced exploration strategies could incorporate additional heuristics such as program structure analysis, dependency tracking, or feedback from previous analysis runs to further improve the efficiency of vulnerability discovery.

9

Usage of AI

For the development of this thesis, AI-assisted technologies, specifically large language models, were utilized to enhance various aspects of the writing and research process.

- **Text Transformation and Fluency:** AI tools were primarily used to refine and transform sections of the text to improve fluency, clarity, and highlight important aspects without altering the original content or technical accuracy. This included rephrasing sentences, improving sentence structure, and ensuring a consistent academic tone.
- **Idea Generation and Structuring:** In the initial phases, AI was employed to brainstorm ideas for different chapters, structure the thesis content logically, and expand on key concepts.
- **Grammar and Spelling Checks:** AI tools assisted in reviewing the thesis for grammatical errors, spelling mistakes, and punctuation issues.

Bibliography

- [1] Thanassis Avgerinos, Alexandre Rebert, Sang Kil Cha, and David Brumley. Enhancing symbolic execution with veritesting. In *Proceedings of the 36th International Conference on Software Engineering*, ICSE 2014, pages 1083–1094. ACM, 2014.
- [2] Volodymyr Kuznetsov, Johannes Kinder, Stefan Bucur, and George Candea. Efficient state merging in symbolic execution. In *Proceedings of the 33rd ACM SIGPLAN Conference on Programming Language Design and Implementation*, PLDI '12, pages 193–204. ACM, 2012.
- [3] Jiang Ming, Dinghao Wu, Jun Wang, Xinyu Xing, and Zhiqiang Liu. TaintPipe: Pipelined symbolic taint analysis. In *24th USENIX Security Symposium*, pages 65–80. USENIX Association, 2015.
- [4] Saahil Ognawala, Martín Ochoa, Alexander Pretschner, and Tobias Limmer. MACKE: Compositional analysis of low-level vulnerabilities with symbolic execution. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, SAC '16, pages 1623–1628. ACM, 2016.
- [5] Sebastian Poeplau and Aurélien Francillon. Symbolic execution with SymCC: Don't interpret, compile! In *29th USENIX Security Symposium*, pages 181–198. USENIX Association, 2020.
- [6] Edward J. Schwartz, Thanassis Avgerinos, and David Brumley. All you ever wanted to know about dynamic taint analysis and forward symbolic execution (but might have been afraid to ask). *Proceedings of the 2010 IEEE Symposium on Security and Privacy*, pages 317–331, 2010.
- [7] Yan Shoshitaishvili, Ruoyu Wang, Christopher Salls, Nick Stephens, Mario Polino, Audrey Dutcher, John Grosen, Siji Feng, Christophe Hauser, Christopher Kruegel, and Giovanni Vigna. SoK: (state of) the art of war: Offensive techniques in binary analysis. In *IEEE Symposium on Security and Privacy*, pages 138–153. IEEE, 2016.
- [8] Jake Springer and Siji Feng. Teaching with angr: A symbolic execution curriculum and CTF. In *2018 IEEE/ACM 1st International Workshop on Automated Software Engineering Education*, pages 13–20. IEEE, 2018.
- [9] Shuangjie Yao and Junjie Chen. Empc: Effective path prioritization for symbolic execution with path cover. *arXiv preprint arXiv:2505.03555*, 2025. Available at: <https://arxiv.org/abs/2505.03555>.

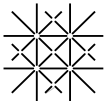


Appendix

A.1 State Explosion Test Program Source Code

```
1 #include <stdio.h>
2 #include <string.h>
3
4 void create_many_string_states(char *input) {
5     // Create branching based on string content (tainted)
6     if (strlen(input) > 10) {
7         printf("Long input\n");
8     } else {
9         printf("Short input\n");
10    }
11
12    if (input[0] == 'A') {
13        printf("Starts with A\n");
14    } else if (input[0] == 'B') {
15        printf("Starts with B\n");
16    } else {
17        printf("Starts with other\n");
18    }
19 }
20
21 void untainted_integer_branching() {
22     int static_val = 25; // Hardcoded, not tainted
23
24     // Classical angr will explore these, TraceGuard should skip
25     if (static_val > 20) {
26         printf("Static val > 20\n");
27     }
28     if (static_val % 5 == 0) {
29         printf("Static val divisible by 5\n");
30     }
31     if (static_val < 30) {
```

```
32     printf("Static val < 30\n");
33 }
34 }
35
36 void critical_vulnerability(char *dangerous_input) {
37     char tiny_buffer[6]; // Extremely small
38     strcpy(tiny_buffer, dangerous_input); // Obvious overflow
39     printf("Critical: %s\n", tiny_buffer);
40 }
41
42 int main() {
43     char user_input[400];
44
45     // Untainted branching - TraceGuard should skip
46     untainted_integer_branching();
47
48     printf("Enter string input: ");
49     if (fgets(user_input, sizeof(user_input), stdin)) {
50         // Remove newline
51         user_input[strcspn(user_input, "\n")] = 0;
52
53         // This creates tainted string-based branching
54         create_many_string_states(user_input);
55
56         // Clear vulnerability with tainted data
57         critical_vulnerability(user_input);
58     }
59     return 0;
60 }
```



Declaration on Scientific Integrity

(including a Declaration on Plagiarism and Fraud)

Translation from German original

Title of Thesis: Optimizing Symbolic Execution Through Taint Analysis and Path Prioritization

Name Assessor: Dr. Marco Vogt


Name Student: Ruben Hutter

Matriculation No.: 2020-065-934

I attest with my signature that I have written this work independently and without outside help. I also attest that the information concerning the sources used in this work is true and complete in every respect. All sources that have been quoted or paraphrased have been marked accordingly.

Additionally, I affirm that any text passages written with the help of AI-supported technology are marked as such, including a reference to the AI-supported program used.

This paper may be checked for plagiarism and use of AI-supported technology using the appropriate software. I understand that unethical conduct may lead to a grade of 1 or "fail" or expulsion from the study program.

Place, Date: 02.07.2025 Student: 

Will this work, or parts of it, be published?


☐

No



Yes. With my signature I confirm that I agree to a publication of the work (print/digital) in the library, on the research database of the University of Basel and/or on the document server of the department. Likewise, I agree to the bibliographic reference in the catalog SLSP (Swiss Library Service Platform). (cross out as applicable)

Publication as of: 03.07.2025

Place, Date: 02.07.2025 Student: 

Place, Date: _____ Assessor: _____

Please enclose a completed and signed copy of this declaration in your Bachelor's or Master's thesis.