Security Anomalies and Continuous Vulnerability Detection

Luciano Sampaio - lsampaio@inf.puc-rio.br
Alessandro Garcia - afgarcia@inf.puc-rio.br

LES | DI |PUC-Rio - Brazil

OPUS Research Group

Introduction

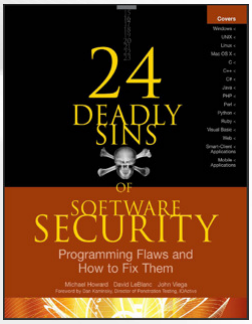
- Secure programming is the practice of writing programs that are resistant to attacks by malicious people or programs

- Security vulnerability or just vulnerability is a flaw that can be exploited to allow an attacker to cause unintended operations

- 86% of audited websites has at least 1 serious security vulnerability (WhiteHat - 2013)

Top 10 vulnerabilities

- Open Web Application Security Project (OWASP)
  - 2003, 2004, 2007, 2010, 2013
- OWASP Top 10 – 2013
  - From 8 datasets from 7 companies with over 500,000 vulnerabilities
  - 01 – (SQL/Command) Injection
  - 02 – Broken Authentication and Session Management
  - 03 – Cross-Site Scripting (XSS)
  - 04 – Insecure Direct Object References
  - 05 – Security Misconfiguration
  - 06 – Sensitive Data Exposure
  - 07 – Missing Function Level Access Control
  - 08 – Cross-Site Request Forgery (CSRF)
  - 09 – Using Known Vulnerable Components
  - 10 – Unvalidated Redirects and Forwards

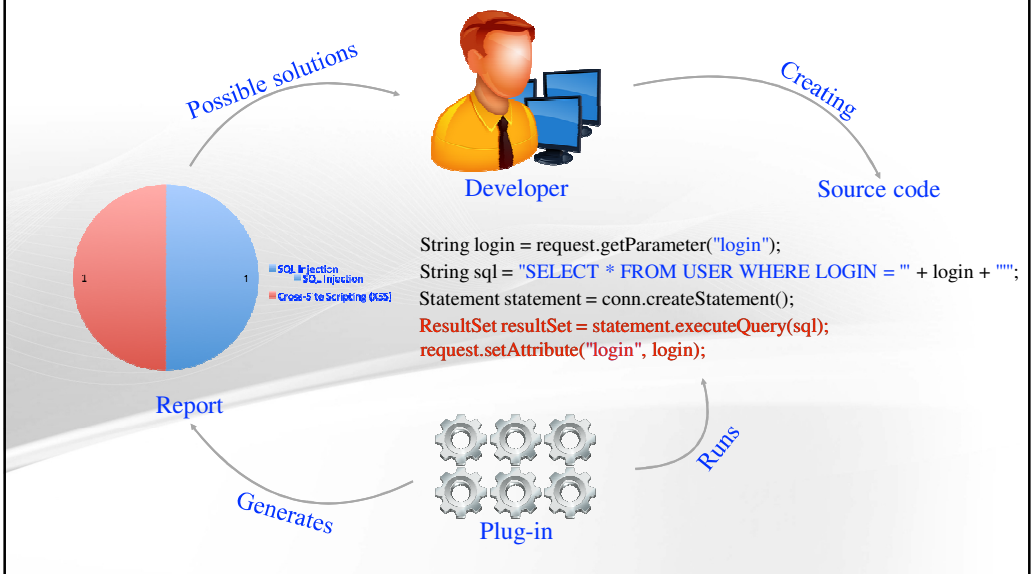


Static analysis can help

- SSVChecker, FindBug, ASIDE, Lapse+, CodePro Analytics, CodeProfiler, JeSS and AppScan IBM
- Key characteristics
  - Late detection
  - Pattern matching

How is usually done… Late detection

Early detection can help



Early detection requirements

- High accurate detection technique
  - Rate of false positives
- The usage of time and resources can not disturb the developer

Early detection using low accurate technique

```
17⊖  @Override
17⊖    @Override
17⊖    @Override
18     protected void doGet(HttpServletRequest request,
19         HttpServletResponse response)
20             throws ServletException, IOException {
21       request.getParameter("")
22     }
24     }
```

Frequently used technique

- Pattern matching is a technique for checking if a pattern matches a given sequence of tokens (letters, numbers, punctuation, and certain symbols)
  - 20/30% of false positives - (Nadeem 2012)

Pattern matching example

```
16  @Override
17  protected void doGet(HttpServletRequest request,
18      HttpServletResponse response)
19          throws ServletException, IOException {
20    PrintWriter printWriter = response.getWriter();
21
22    printWriter.print("a");
23    printWriter.print(("b"));
24
25    String d = "d";
26    printWriter.print((null != "") ? "c" : d);
27
28    printWriter.print(getContent(request));
29    printWriter.print(Boolean.parseBoolean(request.getParameter("bad")));
30  }
31
32  private String getContent(HttpServletRequest request) {
33    int i = 5;
34    if (i > 10) {
35      return request.getParameter("bad");
36    } else if (i == 5) {
37      String bad = request.getParameter("bad");
38
39      return bad;
40    }
41
42    return "ok";
43  }
```

Pattern Matching technique

Data-flow analysis as an alternative

- Data-flow analysis (DFA) is a technique for gathering information about the possible set of values calculated at various points in a computer program
- Originally created and commonly used for implementing optimizations on compilers
- Currently, there is only one solution that uses DFA to detect security vulnerabilities (CodePro)

Data-flow analysis example

Key limitations of the state-of-the-art

- Late detection does not support secure programming but rather security analysis
- Frequently used vulnerability detection techniques present a high rate of false positives
  - False positives are even more critical in early detection
  - Pattern matching or primitive DFA

Research questions

- 01 - Can advanced DFA decrease the rate of false positives when compared to other techniques ?
- 02 - Can the early detection approach help developers produce more secure code when compared to late detection ?

Proposed solutions

- Propose to support a change from the default behavior of late detection to early detection
- Propose new heuristics using a technique named context-sensitive data flow analysis
  - Pattern matching
  - Context-insensitive data flow analysis
- Designed and implemented a prototype
- Performed 2 empirical studies

How to identify something as unsafe/safe ?

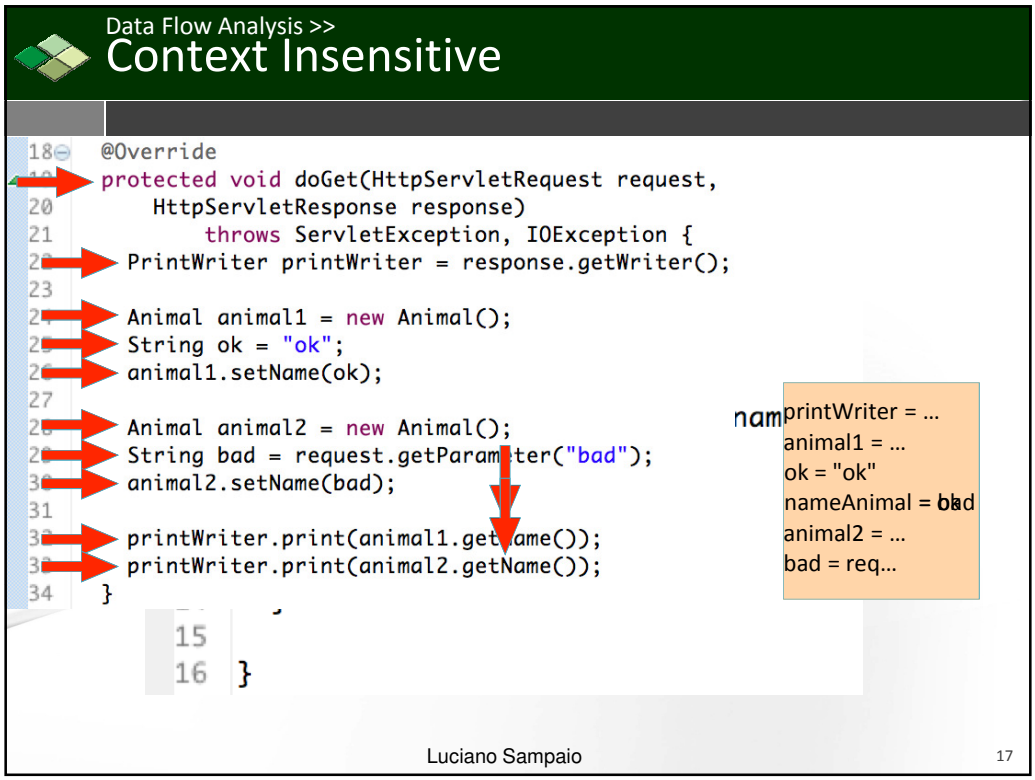
- Entry-Point
- Sanitization-Point
- Exit-Point

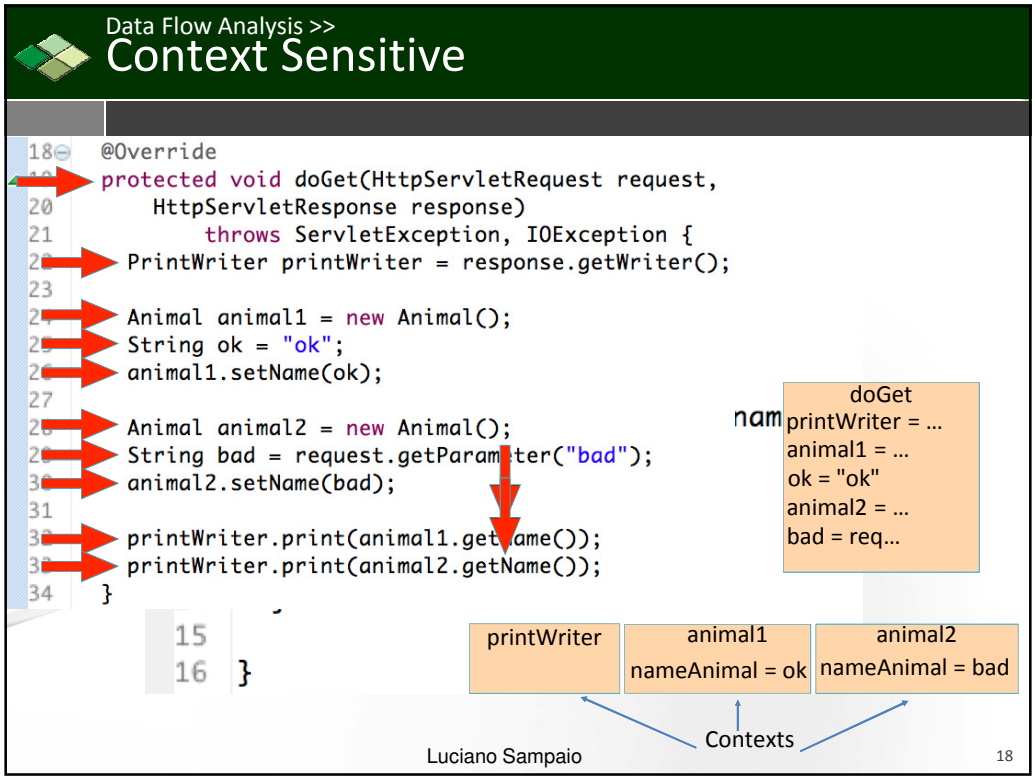
```
17  @Override
18  protected void doGet(HttpServletRequest request,
19      HttpServletResponse response)
20  throws ServletException, IOException {
21
22    String unsafeLogin = request.getParameter("login");
23
24    String safeLogin = ESAPI.encoder().encodeForHTML(unsafeLogin);
25
26    request.setAttribute("login", safeLogin);
27
28  }
```

Entry-Point

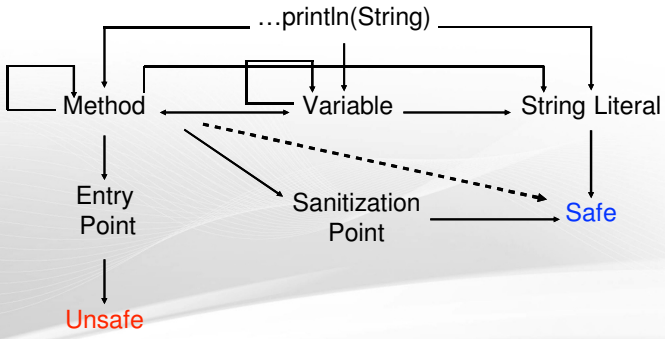
Sanitization-Point

Exit-Point


Data Flow Analysis >>
Context Insensitive
@Override
protected void doGet(HttpServletRequest request,
HttpServletResponse response)
throws ServletException, IOException {
PrintWriter printWriter = response.getWriter();
Animal animal1 = new Animal();
String ok = "ok";
animal1.setName(ok);
Animal animal2 = new Animal();
String bad = request.getParameter("bad");
animal2.setName(bad);
printWriter.print(animal1.getName());
printWriter.print(animal2.getName());
}
15
16 }
printWriter = ...
animal1 = ...
ok = "ok"
animal2 = ...
bad = req...
Luciano Sampaio
17


Data Flow Analysis >>
Context Sensitive
@Override
protected void doGet(HttpServletRequest request,
HttpServletResponse response)
throws ServletException, IOException {
PrintWriter printWriter = response.getWriter();
Animal animal1 = new Animal();
String ok = "ok";
animal1.setName(ok);
Animal animal2 = new Animal();
String bad = request.getParameter("bad");
animal2.setName(bad);
printWriter.print(animal1.getName());
printWriter.print(animal2.getName());
}
15
16 }
doGet
printWriter = ...
animal1 = ...
ok = "ok"
animal2 = ...
bad = req...
printWriter
animal1
nameAnimal = ok
animal2
nameAnimal = bad
Contexts
Luciano Sampaio
18

Our proposed heuristics



Data Flow Analysis >>

Supported vulnerabilities

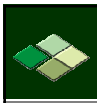
- Vulnerabilities that stem from program input and output not being properly validated are recognized as being the most common ones
- These vulnerabilities are not dependent on how they are implemented

- 01 - Command Injection
- 02 - Cookie Poisoning
- 03 - Cross-Site Scripting (XSS)
- 04 - HTTP Response Splitting
- 05 - LDAP Injection
- 06 - Log Forging
- 07 - Path Traversal
- 08 - Reflection Injection
- 09 - Security Misconfiguration
- 10 - SQL Injection
- 11 - XPath Injection

Data Flow Analysis >>

New type of problems - Infinite Loop

```
16  @Override
17  protected void doGet(HttpServletRequest request,
18      HttpServletResponse response)
19          throws ServletException, IOException {
20    infiniteLoop(request, response);
21  }
22
23  private void infiniteLoop(HttpServletRequest request,
24      HttpServletResponse response)
25          throws IOException {
26    PrintWriter printWriter = response.getWriter();
27
28    String a = request.getParameter("a");
29    printWriter.print(a);
30
31    infiniteLoop(request, response);
32
33    String b = request.getParameter("b");
34    printWriter.print(b);
35  }
```

Data Flow Analysis >>

Current limitations of our implementation

- Containers
- InnerClasses

```
13  @Override
14  protected void doGet(HttpServletRequest request,
15      HttpServletResponse response)
16          throws ServletException, IOException {
17    String a = request.getParameter("a");
18    String b = request.getParameter("b");
19
20    String[] x = { a, b, "c" };
21
22    response.getWriter().print(x);
23    response.getWriter().print(x[0]);
24    response.getWriter().print(x[1]);
25    response.getWriter().print(x[2]);
26    response.getWriter().print(x[50]);
27  }
```

ESVD

Dillig, I., Dillig, T. and Aiken, A. (2011). Precise reasoning for programs using containers. ACM SIGPLAN Notices

Evaluation

- Rate of false positives
  - Exploratory study - Benchmark on 5 open-source projects and 1 custom-made project
- Early detection effectiveness
  - Controlled experiment - Participants were asked to create a code using our tool

Evaluation >>

Study 1: Accuracy Benchmarking

	Blueblog	Personalblog	WebGoat	Roller	Pebble	NCO
Version	1.0	1.2.6	5.4	0.9.9	2.6.4	1.0
Number of packages	22	10	24	70	100	49
Number of classes	38	38	159	283	743	84
Number of methods	227	253	1.453	2.704	3.445	517
Lines of Code	2.200	2.933	24.483	34.301	36.709	6.048
Number of Vulnerabilities	18	148	488	521	440	77

Analyzed projects

	Pattern Matching	Data Flow Analysis - CI	Data Flow Analysis - CS
Lapse+	X		
ASIDE	X		
CodePro		X	
ESVD			X

Selected solutions

Study 1 >>
Analyzed vulnerabilities

		Pattern Matching		DFA - CI	DFA - CS
Nr	Vulnerability	ASIDE	Lapse+	CodePro	ESVD
1	**Command Injection**	0	1	1	1
2	**Cookie Poisoning**	1	1	1	1
3	**Cross-Site Scripting (XSS)**	1	1	1	1
4	**HTTP Response Splitting**	0	1	1	1
5	**LDAP Injection**	0	1	1	1
6	**Log Forging**	1	1	1	1
7	**Path Traversal**	0	1	1	1
8	**Reflection Injection**	0	0	1	1
9	**Security Misconfiguration**	0	0	1	1
10	**SQL Injection**	1	1	1	1
11	**XPath Injection**	0	1	1	1
	Total	**4**	**9**	**11**	**11**

Study 1 >>
Summary

	Precision	Recall	F1 Score	% False Positive
ASIDE	0,48	0,39	0,43	51,78%
CodePro	0,62	0,07	0,13	37,62%
Lapse+	0,55	0,36	0,43	44,73%
ESVD	0,88	0,66	0,75	11,70%

Study 1 >>

Summary of False Positives

	Blueblog	Personalblog	WebGoat	Roller	Pebble	NCO
ASIDE	74%	13%	51%	70%	50%	29%
Lapse+	20%	25%	50%	22%	29%	64%
CodePro	59%	17%	32%	59%	54%	61%
ESVD	0%	3%	21%	1%	5%	61%

Results of false positives per analyzed project

Study 1 >>

WebGoat - False Positive

```
272     int nextId = getNextUID(s);
273     String query = "INSERT INTO employee VALUES ( " + nextId + ", ?,?,?,?,?,?,?,?,?,?,?,?,?)";
274
275     // System.out.println("Query: " + query);
276
277     try
278     {
279       PreparedStatement ps = WebSession.getConnection(s).prepareStatement(query);
280
281       ps.setString(1, employee.getFirstName().toLowerCase());
282       ps.setString(2, employee.getLastName());
283       ps.setString(3, employee.getSsn());
284       ps.setString(4, employee.getTitle());
285       ps.setString(5, employee.getPhoneNumber());
286       ps.setString(6, employee.getAddress1());
287       ps.setString(7, employee.getAddress2());
288       ps.setInt(8, employee.getManager());
289       ps.setString(9, employee.getStartDate());
290       ps.setString(10, employee.getCcn());
291       ps.setInt(11, employee.getCcnLimit());
292       ps.setString(12, employee.getDisciplinaryActionDate());
293       ps.setString(13, employee.getDisciplinaryActionNotes());
294       ps.setString(14, employee.getPersonalDescription());
295
296       ps.execute();
```

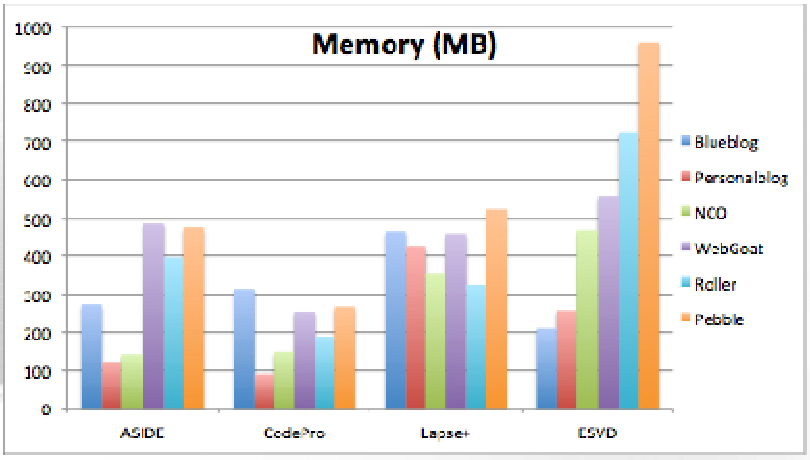
Description	Line	Vulnerability	Resource	Path
query has 1 vulnerable path.	279	Sql Injection	UpdateProfile.java	
String concatenation is not allowed on queries.	273	String concatenation	UpdateProfile.java	handleRequest - this.createEmployeeProfile(s,userId,

Study 1 >>
NCO - False Positive

```
384   private static final String TABLE_NAME    = "APPOINTMENT-";
385   private static final String SELECT_BY_ID  = TABLE_NAME + "SELECT_BY_ID";
386
387   public Appointment selectById(Appointment entity) throws DAOException {
388     // The objects that will be used in this method.
389     Connection conn = null;
390     PreparedStatement pst = null;
391     ResultSet rs = null;
392
393     try {
394       // It opens the connection.
395       conn = getConnection();
396
397       // It gets the sql from the Map Query file.
398       String query = getMapQuery().get(SELECT_BY_ID);
399
400       // It creates a prepared statement to interact with the database.
401       pst = conn.prepareStatement(query);
402       pst.setInt(1, entity.getId());
403       getLogger().debug(pst.toString());
```

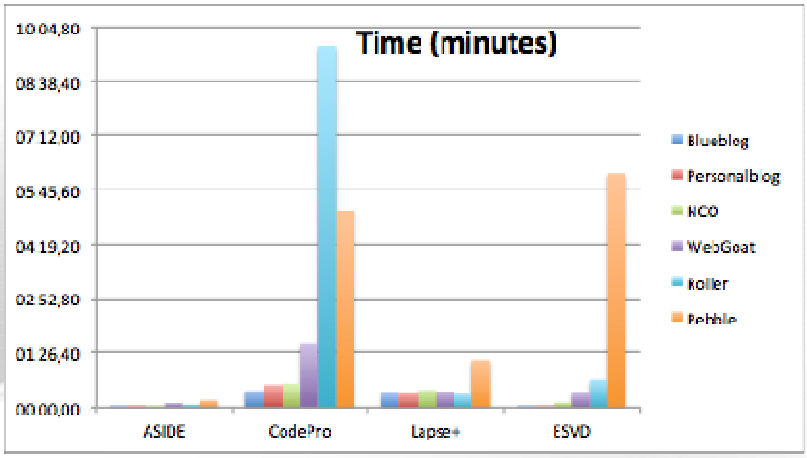
Description	Line	Vulnerability	Resource	Path
query has 3 vulnerable paths.	401	Sql Injection	AppointmentDAO.java	
String concatenation is not allowed on queries.	385	String concatenation	AppointmentDAO.java	details - bll.selectById(getEntity
String concatenation is not allowed on queries.	385	String concatenation	AppointmentDAO.java	editAppointmentConfirmed - ed
String concatenation is not allowed on queries.	385	String concatenation	AppointmentDAO.java	editConfirmed - edit() - bll.sele

Study 1 >>
Memory



The applications are ordered by size

Study 1 >>
Time (Minutes)



The applications are ordered by size

Study 1 >>
Findings

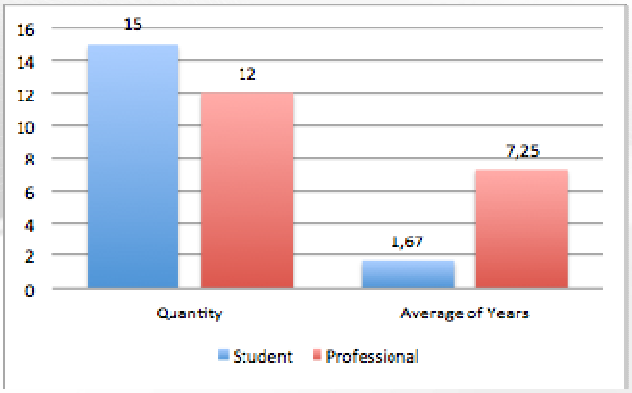
- We achieved 11,70% of rate of false positives
  - The best pattern-matching result was 44,73%
- There is a trade-off, better results mean more time and memory usage
  - This can be problem for large projects when using DFA-CS
- RQ1 - Can DFA-CS decrease the rate of false positives when compared to other techniques ? YES!

Evaluation >>
Study 2: Late vs. Early Detection – A Quasi-Experiment

- 2 groups of participants (students and professionals), divided in 2 groups (Early Detection and Late Detection)
  - Both using ESVD
- Asked them to develop some functionalities of a small system
  - Initial project and basic jsp pages already created
  - Login and logout
  - Add, Update, Delete and List comments
- Recorded their screen, audio and Eclipse's interactions
  - ScreenFlow
  - Rabbit-eclipse

Study 2 >>
Participants

	Early	Late	Total	Total
Student	10	10	20	34
Professional	7	7	14	



Study 2 >>
Participants - Final numbers

	Early	Late	Total	Total
Student	10	10	20	34
Professional	7	7	14	

	Early	Late	Total	Total
Student	2	6	8	18
Professional	6	4	10	

Study 2 >>
Programming timing and completed tasks

	Programming time		
	Early	Late	
Professional	9:32:40	4:31:07	**14:03:47**
Student	1:44:18	2:46:15	**4:30:33**
Sum	**11:16:58**	**7:17:22**	
Total	**18:34:20**		

Task 1	Task 2	Task 3	Task 4	Task 5
18	**8**	**4**	**2**	**2**

Study 2 >>
Number of vulnerabilities

	Added		Removed		Left	
	Early	Late	Early	Late	Early	Late
Professional	31	9	10	1	21	8
Student	4	13	2	1	2	12
	57		**14**		**43**	

Vulnerability	Added	Removed	Left
HTTP Response Splitting	1	1	0
Cookie Poisoning	2	0	2
SQL Injection	3	1	2
Log forging	10	6	4
Cross-Site Scripting	14	3	11
Misconfiguration	27	3	24
Total	**57**	**14**	**43**

Study 2 >>
Findings

- During the experiment, 57 security vulnerabilities were added
  - Early detection group added 35 vulnerabilities and removed 12 (or 34,2%) vulnerabilities
  - Late detection group added 22 vulnerabilities and only removed 2 (or 9,09%)
- RQ2 - Can the early detection approach help developers produce more secure code when compared to late detection ? **YES!**

Conclusion

- Based on our two studies:
  - Data flow analysis with context-sensitivity reduced the rate of false positives when compared to other techniques
  - Early detection combined with DFA-CS helped developers produce more secure code

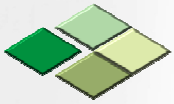
Conclusions

- The heuristic strategies capable of finding 11 security vulnerabilities that stem from input and output not being properly sanitized
- Proposal and implementation of the algorithm of data flow analysis with context sensitivity to find security vulnerabilities
- The complete list with known security vulnerabilities (ground truth) for each of the analyzed open-source projects

Future work

- Increase the number of supported vulnerabilities
  - We currently support 11 types
- Add a ranking system for the found vulnerabilities
  - Asked by several participants
- Allow developers to add, edit or remove methods from the lists of *entry-points*, *exit-points* and *sanitization-points*

Additional slides



OPUS Research Group

Entry-Point

- An *entry-point,* also referred as source, is a point in the source code where external and untrusted input enters the application
- We have 81 entry-points registered

```
<entrypoint id="01">
  <qualifiedname>javax.servlet.ServletRequest</qualifiedname>
  <methodname>getAttribute</methodname>
  <parameters type="java.lang.String" />
</entrypoint>
<entrypoint id="02">
  <qualifiedname>javax.servlet.ServletRequest</qualifiedname>
  <methodname>getAttributeNames</methodname>
</entrypoint>
<entrypoint id="03">
  <qualifiedname>javax.servlet.ServletRequest</qualifiedname>
  <methodname>getCharacterEncoding</methodname>
</entrypoint>
```

Luciano Sampaio 43

Sanitization-Point

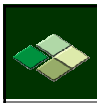
- A *sanitization-point,* also referred as sanitizer, is a point in the source code where a method or class receives an untrusted input and returns it as a trusted output
- We have 52 sanitization-points registered

```
<sanitizer id="01">
  <qualifiedname>org.owasp.encoder.Encode</qualifiedname>
  <methodname>forHtml</methodname>
  <parameters type="java.lang.String" />
</sanitizer>
<sanitizer id="02">
  <qualifiedname>org.owasp.encoder.Encode</qualifiedname>
  <methodname>forHtmlContent</methodname>
  <parameters type="java.lang.String" />
</sanitizer>
<sanitizer id="03">
  <qualifiedname>org.owasp.encoder.Encode</qualifiedname>
  <methodname>forHtmlAttribute</methodname>
  <parameters type="java.lang.String" />
</sanitizer>
```

Luciano Sampaio 44

Exit-Points >>
Accepted Rules

- 0 - Anything
- 1 - Sanitized
- 2 - Null
- 4 - Literal
- 8 - Concatenation (used for SQL Injection)
- …



Pattern Matching vs Data Flow Analysis

Pattern Matching

Data Flow Analysis

Data Flow Analysis >>
Early Detection

```
@Override
protected void doGet(HttpServletRequest request,
    HttpServletResponse response)
        throws ServletException, IOException {
  request.getParameter("")
}
```

```
@Override
protected void doGet(HttpServletRequest request,
    HttpServletResponse response)
        throws ServletException, IOException {
  String bad = request.getParameter("bad");
}
```

```
@Override
protected void doGet(HttpServletRequest request,
    HttpServletResponse response)
        throws ServletException, IOException {
  String bad = request.getParameter("bad");

  response.getWriter().print(bad);
}
```

```
@Override
protected void doGet(HttpServletRequest request,
    HttpServletResponse response)
        throws ServletException, IOException {
  String bad = request.getParameter("bad");

  String safe = ESAPI.encoder().encodeForHTML(bad);

  response.getWriter().print(safe);
  response.getWriter().print(ESAPI.encoder().encodeForHTML(bad));
}
```

ESVD

Study 2 >>
Searching for help

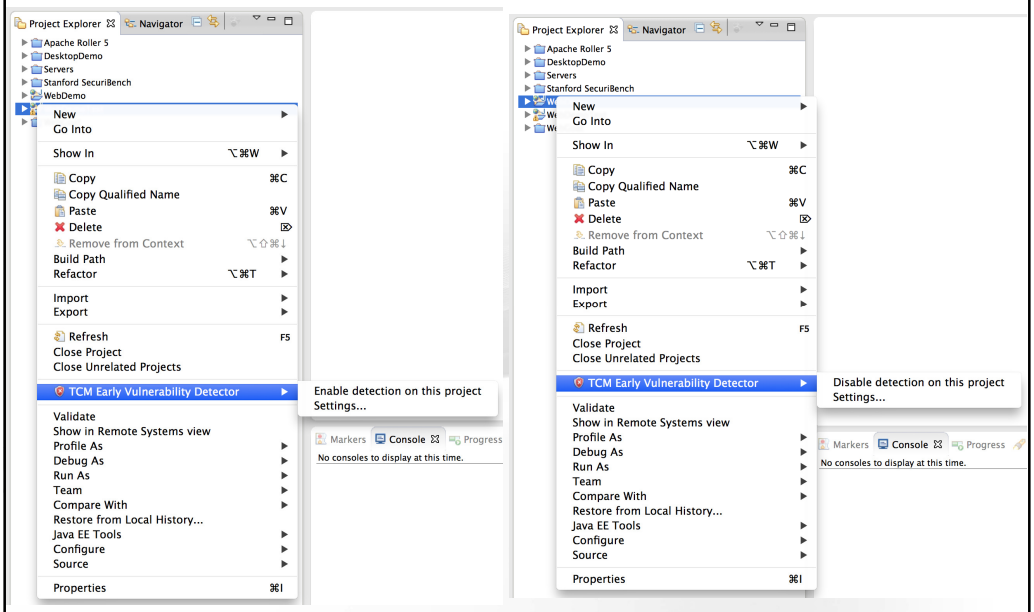
Link	Nr Times
1	27
2	8
3	5
4	6
6	4
7	1
10	1
	52

- You can learn from HelloWold, but should never use its source code.

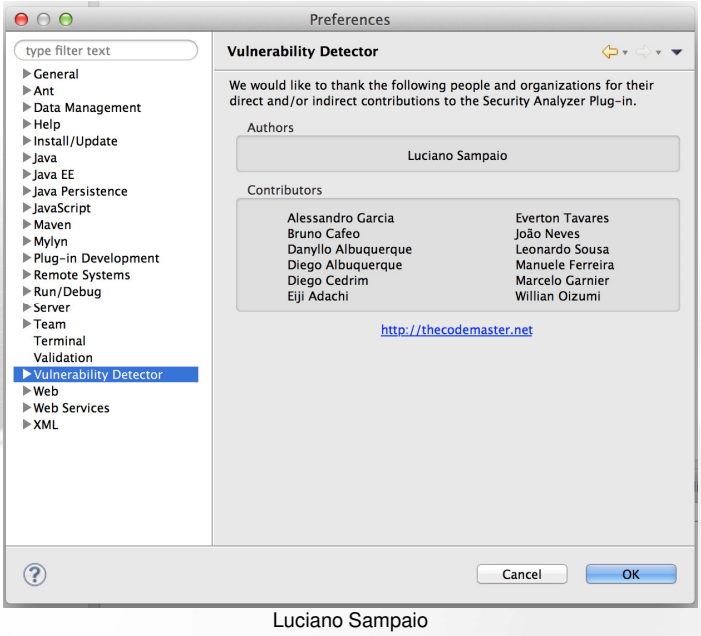
The plug-in

- ESVD - Early Security Vulnerability Detector - 0.3.9;
- Download at: (FREE)
  - https://marketplace.eclipse.org/content/early-security-vulnerability-detector-esvd/
- A project containing several security vulnerabilities:
  - http://www.inf.puc-rio.br/~lsampaio/plugin/early_vulnerability_detector/latest/WebDemo.zip
- How to use ESVD: (Portuguese)
  - https://www.youtube.com/watch?v=pNr38gMWvHQ
- More info at: http://thecodemaster.net/

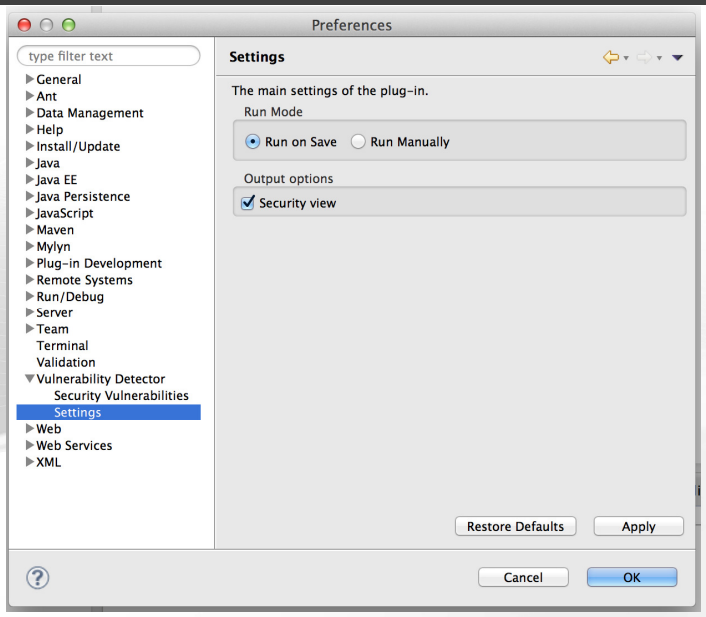
The plug-in >> Menu

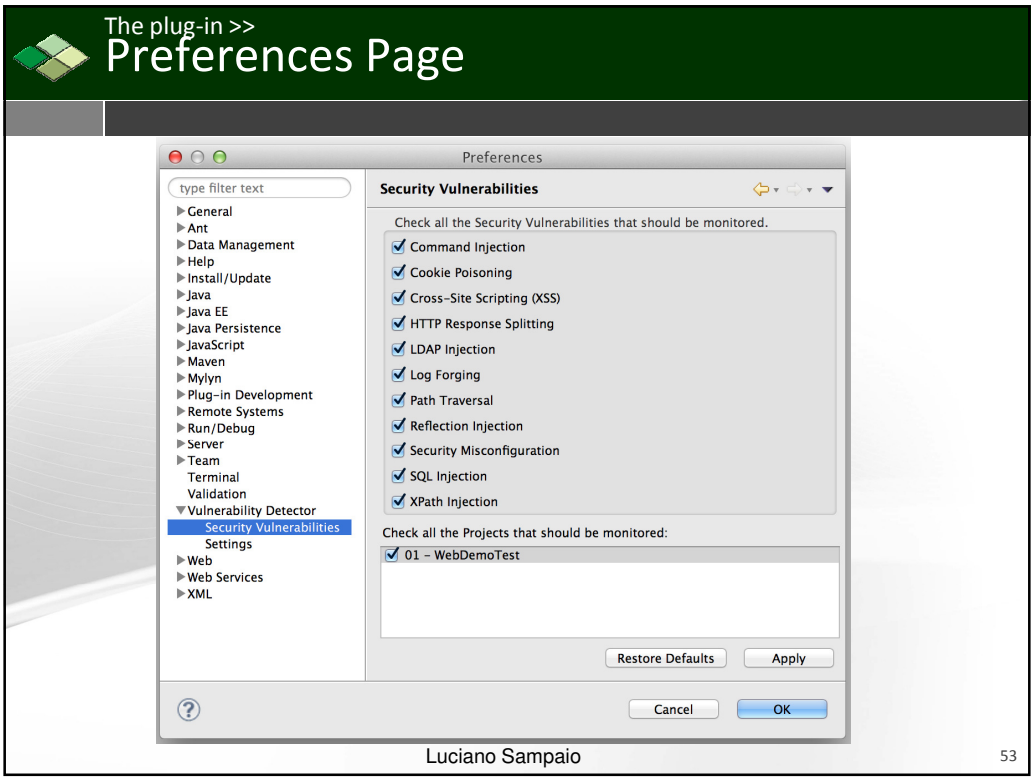


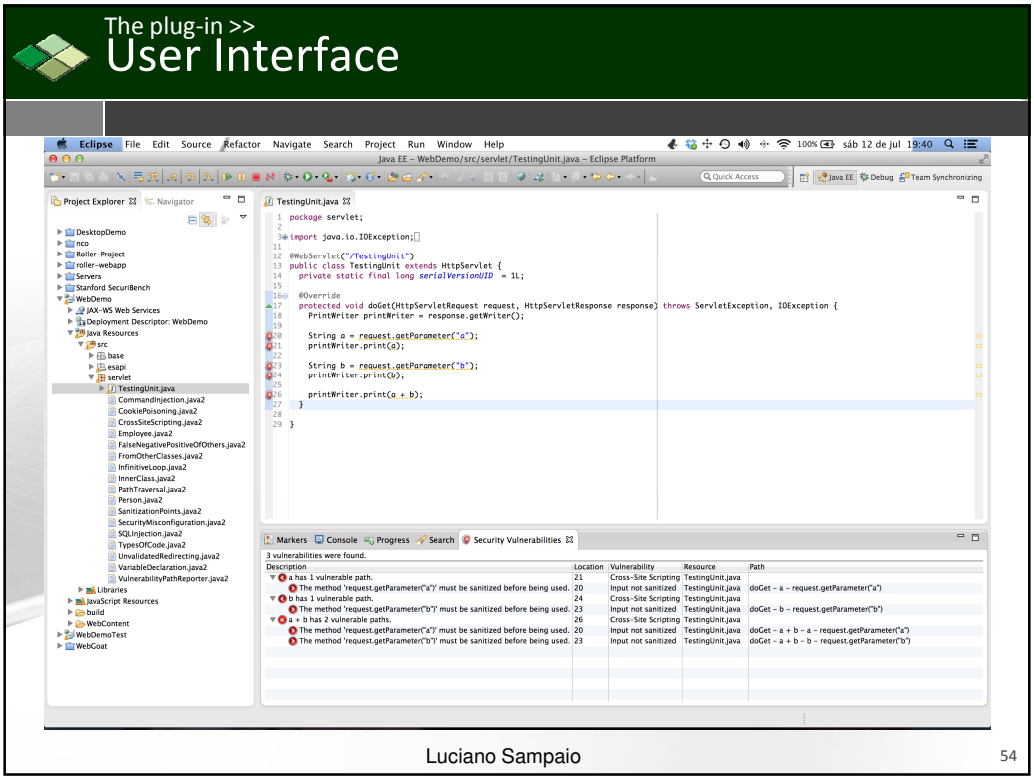
The plug-in >> Preferences Page

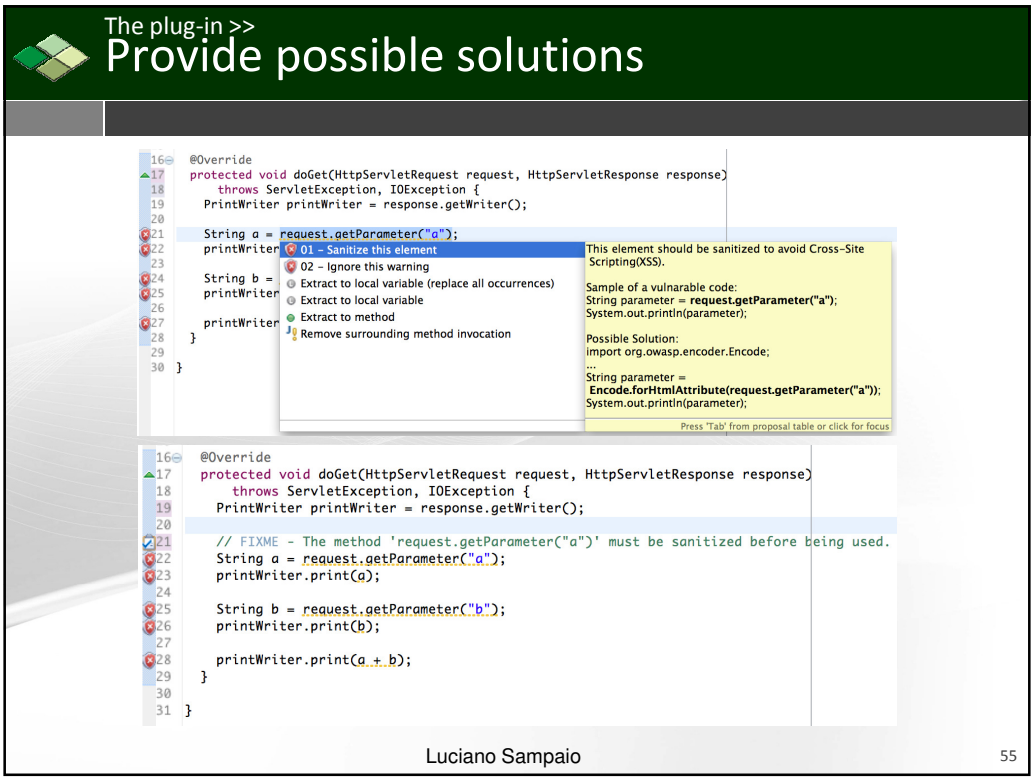


The plug-in >> Preferences Page




The plug-in >>
Preferences Page
Preferences
type filter text
General
Ant
Data Management
Help
Install/Update
Java
Java EE
Java Persistence
JavaScript
Maven
Mylyn
Plug-in Development
Remote Systems
Run/Debug
Server
Team
Terminal
Validation
Vulnerability Detector
Security Vulnerabilities
Settings
Web
Web Services
XML
Security Vulnerabilities
Check all the Security Vulnerabilities that should be monitored.
Command Injection
Cookie Poisoning
Cross-Site Scripting (XSS)
HTTP Response Splitting
LDAP Injection
Log Forging
Path Traversal
Reflection Injection
Security Misconfiguration
SQL Injection
XPath Injection
Check all the Projects that should be monitored:
01 - WebDemoTest
Restore Defaults
Apply
Cancel
OK
Luciano Sampaio
53


The plug-in >>
User Interface
Eclipse File Edit Source Refactor Navigate Search Project Run Window Help
Java EE - WebDemo/src/servlet/TestingUnit.java - Eclipse Platform
Project Explorer
Navigator
TestingUnit.java
package servlet;
import java.io.IOException;
@WebServlet("/TestingUnit")
public class TestingUnit extends HttpServlet {
private static final long serialVersionUID = 1L;
@Override
protected void doGet(HttpServletRequest request, HttpServletResponse response) throws ServletException, IOException {
PrintWriter printWriter = response.getWriter();
String a = request.getParameter("a");
printWriter.print(a);
String b = request.getParameter("b");
printWriter.print(b);
printWriter.print(a + b);
}
}
Markers
Console
Progress
Search
Security Vulnerabilities
3 vulnerabilities were found.
Description
Location
Vulnerability
Resource
Path
a has 1 vulnerable path.
The method 'request.getParameter("a")' must be sanitized before being used.
b has 1 vulnerable path.
The method 'request.getParameter("b")' must be sanitized before being used.
a + b has 2 vulnerable paths.
The method 'request.getParameter("a")' must be sanitized before being used.
The method 'request.getParameter("b")' must be sanitized before being used.
Luciano Sampaio
54



The plug-in >>
Provide possible solutions

```
16  @Override
17  protected void doGet(HttpServletRequest request, HttpServletResponse response)
18      throws ServletException, IOException {
19    PrintWriter printWriter = response.getWriter();
20
21    String a = request.getParameter("a");
22    printWriter
23
24    String b =
25    printWriter
26
27    printWriter
28  }
29
30 }
```

```
01 – Sanitize this element
02 – Ignore this warning
Extract to local variable (replace all occurrences)
Extract to local variable
Extract to method
Remove surrounding method invocation
```

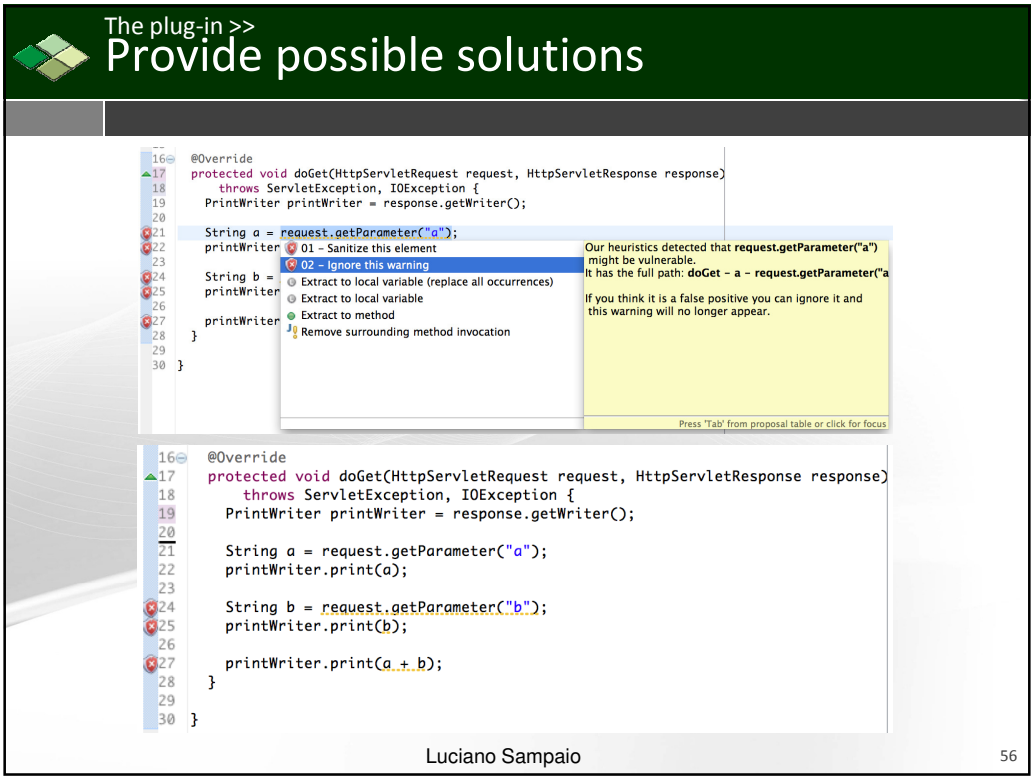
This element should be sanitized to avoid Cross-Site Scripting(XSS).

Sample of a vulnarable code:
String parameter = **request.getParameter("a")**;
System.out.println(parameter);

Possible Solution:
import org.owasp.encoder.Encode;
...
String parameter =
Encode.forHtmlAttribute(request.getParameter("a"));
System.out.println(parameter);

Press 'Tab' from proposal table or click for focus

```
16  @Override
17  protected void doGet(HttpServletRequest request, HttpServletResponse response)
18      throws ServletException, IOException {
19    PrintWriter printWriter = response.getWriter();
20
21    // FIXME - The method 'request.getParameter("a")' must be sanitized before being used.
22    String a = request.getParameter("a");
23    printWriter.print(a);
24
25    String b = request.getParameter("b");
26    printWriter.print(b);
27
28    printWriter.print(a + b);
29  }
30
31 }
```

The plug-in >>
Provide possible solutions

```
16  @Override
17  protected void doGet(HttpServletRequest request, HttpServletResponse response)
18      throws ServletException, IOException {
19    PrintWriter printWriter = response.getWriter();
20
21    String a = request.getParameter("a");
22    printWriter
23
24    String b =
25    printWriter
26
27    printWriter
28  }
29
30 }
```

```
01 – Sanitize this element
02 – Ignore this warning
Extract to local variable (replace all occurrences)
Extract to local variable
Extract to method
Remove surrounding method invocation
```

Our heuristics detected that **request.getParameter("a")** might be vulnerable.
It has the full path: **doGet – a – request.getParameter("a**

If you think it is a false positive you can ignore it and this warning will no longer appear.

Press 'Tab' from proposal table or click for focus

```
16  @Override
17  protected void doGet(HttpServletRequest request, HttpServletResponse response)
18      throws ServletException, IOException {
19    PrintWriter printWriter = response.getWriter();
20
21    String a = request.getParameter("a");
22    printWriter.print(a);
23
24    String b = request.getParameter("b");
25    printWriter.print(b);
26
27    printWriter.print(a + b);
28  }
29
30 }
```

The plug-in

Early Security Vulnerability Detector - ESVD 0.3.9