



UNIVERSITÀ DI TRENTO

Dipartimento di Ingegneria e Scienza dell'Informazione
Department of Information Engineering and Computer Science

Bachelor's Degree in
Computer Science

FINAL DISSERTATION

HTTP SBTL - SECURITY BROWSER-BASED TESTING LANGUAGE AND IMPLEMENTATION

Supervisor

Silvio Ranise

Student

Matteo Bitussi

Co-Supervisors

Andrea Bisegna

Roberto Carbone

Academic year 2021/2022

Ringraziamenti

...thanks to... TODO (in italiano)

Contents

Abstract	3
1 Introduction	5
1.1 Contributions	5
1.2 Structure of the thesis	6
2 Background	7
2.1 Burp Suite community edition	7
2.2 JSON	7
2.3 Regex	7
2.4 IdM protocols: SAML and OAuth	7
2.4.1 OAuth	8
2.4.2 SAML - Security Assertion Markup Language	9
3 Design of testing language	10
3.1 Choosing a language standard	10
3.2 Example: PKCE Downgrade Test	10
3.3 Language structure	12
3.3.1 Test Suite	12
3.3.2 Test	12
3.3.3 Operation	14
3.3.4 Message section	15
3.3.5 Check Object	15
3.3.6 Precondition and Validate Objects	16
3.3.7 Save	16
3.3.8 Message Operation	16
3.3.9 Message Type definition	18
3.4 The oracle	19
3.5 Sessions	20
4 Implementation	21
4.1 General overview of the tool	21
4.2 The tool	21
4.2.1 User Interface of the tool	21
4.2.2 Session managing	23
4.2.3 Test execution	23
4.2.4 Decoding & encoding of parameters	23
4.2.5 SAML certificate managing	23
4.3 Problems and limitations encountered	23
4.3.1 Automation problems	24
4.3.2 Oracle is sometimes ambiguous	24
4.3.3 Interface and user feedbacks	24

5	Use cases	25
5.1	SAML Use-Case	25
5.2	OAuth & OIDC Use-Case	25
6	Related works	26
6.1	Micro-Id-Gym	26
6.2	SSO Testing language and Plugin	26
7	Conclusions and Future Works	27
	Bibliography	27
A	PKCE test Example complete	30
B	Language comparison	31

Abstract

This thesis deals with the work started in my internship at Fondazione Bruno Kessler (FBK) in the context of HTTP browser based protocol testing. In the last years, there was a constant raising number of services that are transitioning from physical to virtual, for instance, health care services and government services. This was done to simplify some procedures, and so, making them faster to be accomplished. But, some concerns arise from this transition, a high quantity of sensitive data could be at risk to unwanted access. Software and security engineers have to implement the service's software, and they have to manage the data that the software is working with. They will rely on browser-based HTTP security protocols to assure the user identity, to secure the connection between the service and the user, and to assure data integrity.

To make sure all these security requirements are met in the implementation of the service, there is the need to test the implemented services against the well-known vulnerabilities of the protocols used. Usually, to do that, a software engineer or a security tester, which will be called tester, is going to act like an attacker that is trying to see, edit, deny the access to the data in the services. A software will intercept the HTTP messages between service and the tester's browser, and it will make possible for the tester to edit or remove the messages, to simulate an attack and see if the service is vulnerable. The tester will have to gather all the well-known vulnerabilities related to the used protocols and test them manually. This is very time-consuming, and could be done improperly if the tester is not qualified to do it. This is why a software that can automatically test the services against the well known-vulnerabilities and that could give a result of whether the services are vulnerable or not is needed. The advantages are that the group of people that could test the implemented services could be enlarged to people which are not qualified to do security testing. To automatically execute these tests, they have to be defined in a format to be then executed by the testing software.

The work in this thesis is about the implementation of the automatic HTTP security testing software and its relative language. Other software have been evaluated before deciding to start this work, some of them missed the automation part, others were limited in terms of the possible test to be done. For these reasons a new software and language has been designed and implemented. Also, during the developing, a test definition for a web security protocol, OAuth has been written in the implemented language.

The implemented software and language have proved to be working in two different use cases, an OAuth test suite, and a SAML testing scenario. The results have been also confirmed by the comparison with the results of similar software. At the end, the objectives fixed at the beginning of the work were accomplished, the software and language created are an useful tool to help the tester to test the services. The definition of an OAuth and SAML test suite comprehending the well-known vulnerabilities is useful as the tester doesn't need to write them in order to test them, just the automation actions has to be rewritten depending on the service to test. There is still some testing to be done to the software, and some small things that could be improved, but overall it is working well.

Glossary

Burp	Burp Suite Community Edition, a web security testing software. 7, 18, 21, 23, 26
OAuth	OAuth, is a SSO protocol. 3, 5–8, 10, 21, 25, 26
OIDC	Open ID Connect, is a SSO protocol. 6, 21, 25, 26
PKCE	Proof Key for Code Exchange, an extension of OAuth. 10, 12, 15, 18, 19
SAML	Security Assertion Markup Language, is an SSO protocol. 7–9, 23
session track	The list of action the browser is told to do. 19, 20, 23, 24, 26

1 Introduction

In the last years, we are seeing a constant transition from physical to virtual, this is the case of bank transactions, government documents, health care data, and almost anything that can be virtualized. This is a great step forward, the storage is optimized and more easily accessible, facilitating the cataloging of the documents. But there are also some big concerns about security of that data, as it is virtual, the access to it is regulated by some authentication protocols instead of a physical identification of the subject accessing the physical documents. This means that it has to be checked whether the person accessing the data is the one that is claiming to be. The three basic principles over which data security is based are confidentiality, integrity and availability of data, this principle is called the CIA triad [6]. If those requirements are not met, sensitive data could be made inaccessible or even at risk of access from unauthorized parties, as a result, critical services for health care could be inaccessible, this is what happened for example in Lazio [14, 7], Italy, on August 2021, where, because of a RansomWare infecting the internal computer network of the region, and keeping it down for multiple days, the covid vaccines and other health-care services could not be booked. Moreover, a lot of data in the systems has been encrypted, and so, denied of any access. It has to be said that preventing RansomWares infections is not the aim of this thesis, but this is a good example of what the consequences of exploiting a vulnerability could be.

In this thesis the considered services are web-services (from now on service) that rely on browser-based HTTP security protocols. Browser-based protocols are the one used to secure web services communications and . The developing has been done with a particular focus over Identity Management (IdM) protocols such as SAML and OAuth but, without limiting the possible applications to these protocols. To ensure that the CIA triad requirements are satisfied, there is a need to test all the implementations of the protocols used by the service, that are liable of the security of the data. The services that provide access to the data has to be designed properly and tested by a software engineer or a security tester (from now on tester), over the well-known vulnerabilities. The well-known vulnerabilities are vulnerabilities of the protocols that have been discovered over time, and that are known to exist, the service should be built in a way to avoid them, and should be tested over them. But, there could arise some problems, for instance, the tester should be qualified to test the protocols and should know how they work, this is not always the case, as software engineers are sometimes responsible for doing that. Another problem is that due to the fast-evolving nature of protocols, new vulnerabilities will eventually be found, and so, new tests will have to be defined, or old tests would have to be edited. This is a lot of work to be done by the tester. The result is that some vulnerabilities could remain undiscovered, representing a weakness in the system. This thesis proposes a new automated testing software and a language to help the tester to test the web-protocols used in the services mentioned above.

Some software to test specific protocols and vulnerabilities already exist [2, 11], but they are usually very specific and hard-coded in a way that the tests could not be easily edited by the tester. Other tools exist, but they are not usually automated. With this work, I wanted to create a software that could be used to test any type of browser-based HTTP protocol and that could specify the tests in an abstract way, allowing the editing and adding. In order to do this, a new language will be used to specify the tests, and a software will be implemented to execute them. The test results will be evaluated by an oracle, that is composed by a series of components that are responsible for telling if the test is passed or not.

In this thesis the design and implementation of this new software will be discussed.

1.1 Contributions

The contributions of this thesis are the following:

- A language to specify tests for browser-based HTTP protocols, that makes possible to customly specify and edit the tests based on the tester needs.
- A methodology to execute the tests specified by the language, having an automation part, to automatically simulate the victim/attacker actions on the browser, and a test execution part, to execute the tests.
- The implementation of a dedicated tool to execute the tests defined by the language.
- Experimental analysis and a test specification of OAuth and OIDC protocols.

1.2 Structure of the thesis

The thesis is structured as follows:

- **Chapter 2, Background:** A brief introduction over the concepts and software used in the rest of the thesis
- **Chapter 3, Design of the testing language:** The design of the tool and the testing language
- **Chapter 4, Implementation:** The implementation of the tool and the testing language
- **Chapter 5, Use cases:** Some examples of application of the software and the language described in this thesis
- **Chapter 6, Related work:** Other works related to the one of this thesis
- **Chapter 7, Conclusions and future works:** Conclusions, results, and works that could be done in the future

2 Background

In this chapter the subjects needed to comprehend the rest of the thesis will be discussed. The most common subjects will be ignored, giving a focus on the more specific and less common ones.

2.1 Burp Suite community edition

Burp Suite Community Edition (from now on Burp) is one of the most used application security testing software for web security testing. It works by the use of a proxy server over which a browser redirects the traffic to. The proxy does like a Man In The Middle attack, taking the input traffic from the browser and replying the HTTP messages (from now on messages) to the target service, giving also transparency over the TLS (Transport Layer Security) or SSL (Secure Socket Layer) encryption. Burp has access to the proxy, it can sniff HTTP packets and can edit them before they are forwarded to the browser or the target service. Burp also gives the possibility of creating custom plugins giving to the developers access to the java API. This is exactly how the plugin (from now on tool) will be implemented, using Burp as a base over which develop the software that will execute the tests. The tool will be able to intercept, read and edit messages that pass through the Burp's proxy by the use of Burp's API.

2.2 JSON

As stated in [4]: “JavaScript Object Notation (JSON) is a text format for the serialization of structured data. It is derived from the object literals of JavaScript. JSON can represent four primitive types (strings, numbers, booleans, and null) and two structured types (objects and arrays). An object is an unordered collection of zero or more name/value pairs, where a name is a string and a value is a string, number, boolean, null, object, or array. An array is an ordered sequence of zero or more values”. In this thesis, the word “name” identifying a name/value pairs, will not be used and will be substituted with “tag”.

2.3 Regex

Regex stands for regular expression, it is a sequence of symbols that define an ensemble of strings that can be matched by it. There is a list of symbols that can be used to specify which characters to match. For example, if the value of the “Host” header in a message has to be found, a regex like this one has to be defined: `Host:\s?.*` This regex will search for the string “Host:” followed by 0 or 1 whitespace, and all the characters that follow until “\n” is found. This is a complete explanation of the symbols used in this example:

- `\s` is the whitespace character
- `?` is an operator that tells that the preceding symbol will be matched 0 or 1 times
- `*` is an operator that tells that the preceding symbol will be matched 0 or more times
- `.` matches any character except line breaks

2.4 IdM protocols: SAML and OAuth

Identity Management (IdM) protocols are protocols that deals with identity management. For the aim of this thesis, SAML (Security Assertion Markup Language) and OAuth 2.0 (from now on OAuth) will be discussed, as they are used in this thesis.

Both SAML and OAuth are Single Sign-On (SSO) protocols, they follow an authentication scheme that allows a user to log in with a single ID and password to any of several related, yet independent, software systems, a more complete explanation can be found in [11].

2.4.1 OAuth

As stated in [12], the OAuth authorization framework enables a third-party application to obtain limited access to an HTTP service, either on behalf of a resource owner by orchestrating an approval interaction between the resource owner and the HTTP service, or by allowing the third-party application to obtain access on its own behalf. A series of messages has to be exchanged between the two parties in order to authenticate a resource owner that wants to access some reserved data in a service. As described in [12], the flow illustrated in Figure 2.1 describes the interaction between the four roles and includes the following steps:

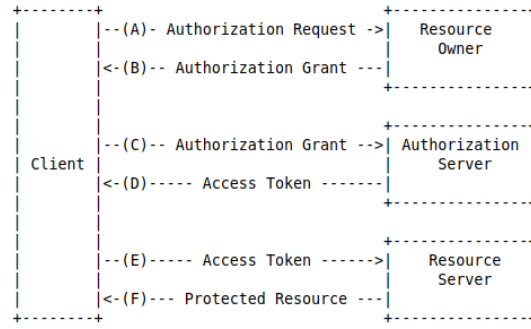


Figure 2.1: OAuth abstract protocol flow source [12]

- (A) The client requests authorization from the resource owner. The authorization request can be made directly to the resource owner (as shown), or preferably indirectly via the authorization server as an intermediary.
- (B) The client receives an authorization grant, which is a credential representing the resource owner’s authorization, expressed using one of four grant types defined in this specification or using an extension grant type. The authorization grant type depends on the method used by the client to request authorization and the types supported by the authorization server.
- (C) The client requests an access token by authenticating with the authorization server and presenting the authorization grant.
- (D) The authorization server authenticates the client and validates the authorization grant, and if valid, issues an access token.
- (E) The client requests the protected resource from the resource server and authenticates by presenting the access token.
- (F) The resource server validates the access token, and if valid, serves the request.

PKCE

As stated in [15]: OAuth 2.0 public clients utilizing the Authorization Code Grant are susceptible to the authorization code interception attack. A technique to mitigate against the threat is the use of Proof Key for Code Exchange (PKCE, pronounced “pixy”). The PKCE extension utilizes a dynamically created cryptographically random key called “`code_verifier`”. A unique `code_verifier` is created for every authorization request, and its transformed value, called “`code_challenge`”, is sent to the authorization server to obtain the authorization code. The authorization code obtained is then sent to the token endpoint with the “`code_verifier`”, and the server compares it with the previously received request code so that it can perform the proof of possession of the “`code_verifier`” by the client. This works as the mitigation since the attacker would not know this one-time key, since it is sent over TLS and cannot be intercepted. The PKCE specification adds additional parameters to the OAuth 2.0 Authorization and Access Token Requests:

- (A) The client creates and records a secret named the “`code_verifier`” and derives a transformed version “`t(code_verifier)`” (referred to as the “`code_challenge`”), which is sent in the OAuth 2.0 Authorization Request along with the transformation method “`t_m`”.

- (B) The Authorization Endpoint responds as usual but records “t(`code_verifier`)” and the transformation method.
- (C) The client then sends the authorization code in the Access Token Request as usual but includes the “`code_verifier`” secret generated at (A).
- (D) The authorization server transforms “`code_verifier`” and compares it to “t(`code_verifier`)” from (B). Access is denied if they are not equal.

An attacker who intercepts the authorization code at (B) is unable to redeem it for an access token, as they are not in possession of the “`code_verifier`” secret.

2.4.2 SAML - Security Assertion Markup Language

As stated in [5], “(SAML) 2.0 is an XML-based framework that allows identity and security information to be shared across security domains. The Assertion, an XML security token, is a fundamental construct of SAML that is often adopted for use in other protocols and specifications. An Assertion is generally issued by an Identity Provider and consumed by a Service Provider that relies on its content to identify the Assertion’s subject for security-related purposes.”

3 Design of testing language

This chapter introduces how the testing software (from now on tool) and the language introduced in Chapter 1 have been designed. The testing language specify what a test does, a test is a series of verifications or modifications done on one or more HTTP messages (from now on messages) exchanged by the client and the service using the browser-based HTTP protocol. After the execution of a test, a check of the expected conditions has to be done, and the result has to be given, to tell if the test has been successfully passed or not. The language will be defined on the basis of all the possible actions which a security tester would need to do on the messages. One of the objectives was to think of a language that could specify the tests once, and that could make possible to test them over multiple web services. For example, a series of tests to verify the well-known vulnerabilities of a specific protocol could be defined and then used on any type of web service.

3.1 Choosing a language standard

A standard to follow while writing the tests had to be found, the first option was to define a proper language with its dedicated parser, this option was discarded, as it was very difficult to implement and other already existing alternatives were available. The one that has been chosen is JSON (JavaScript Object Notation), it allows to specify the hierarchical structure of the tests, using an easy syntax writing. Also, its Object oriented Notation is very useful, as tests and its components will be defined as Objects.

3.2 Example: PKCE Downgrade Test

I want to introduce the language with an example. Due to its complexity, having a real example before the explanation of all its components could be helpful to understand their function. The scenario of this test is based on the presence of a web service hosted in a server (client), the victim, which is using a browser to connect to the client, the Authorization Server (AS) and the tester, that is doing a Man-In-The-Middle attack between the web service and the victim, and between the victim and the AS.

To understand this test, a brief introduction of PKCE has to be done, PKCE (Proof Key for Code Exchange) is an extension of the OAuth protocol, as said in [13]: PKCE(pronounced pixie) extension describes a technique for public clients to mitigate the threat of having the authorization code intercepted. The technique involves the client first creating a secret, and then using that secret again when exchanging the authorization code for an access token. This way if the code is intercepted, it will not be useful, since the token request relies on the initial secret. The PKCE Downgrade test has as objective to test an OAuth vulnerability where removing the parameter `code_challenge` from the url of the authorization request message will be downgrading the authentication process in a way that PKCE will not be used by the AS if the AS is vulnerable [15]. Making possible to the attacker to steal the authorization code from the AS. The message sequence chart of the attack can be seen in Figure 3.1.

To test this, the browser will execute the victim actions, doing a login on the client. The authorization request message has to be intercepted, the `code_challenge` parameter has to be removed from it, and then the message has to be forwarded. The test is passed if the AS does not admit an exchange without PKCE, for instance, it can return an error page, this means that the AS is not vulnerable. Otherwise, if the AS would be vulnerable, the exchange would continue without any error pages or messages. The complete test can be found in Attachment A

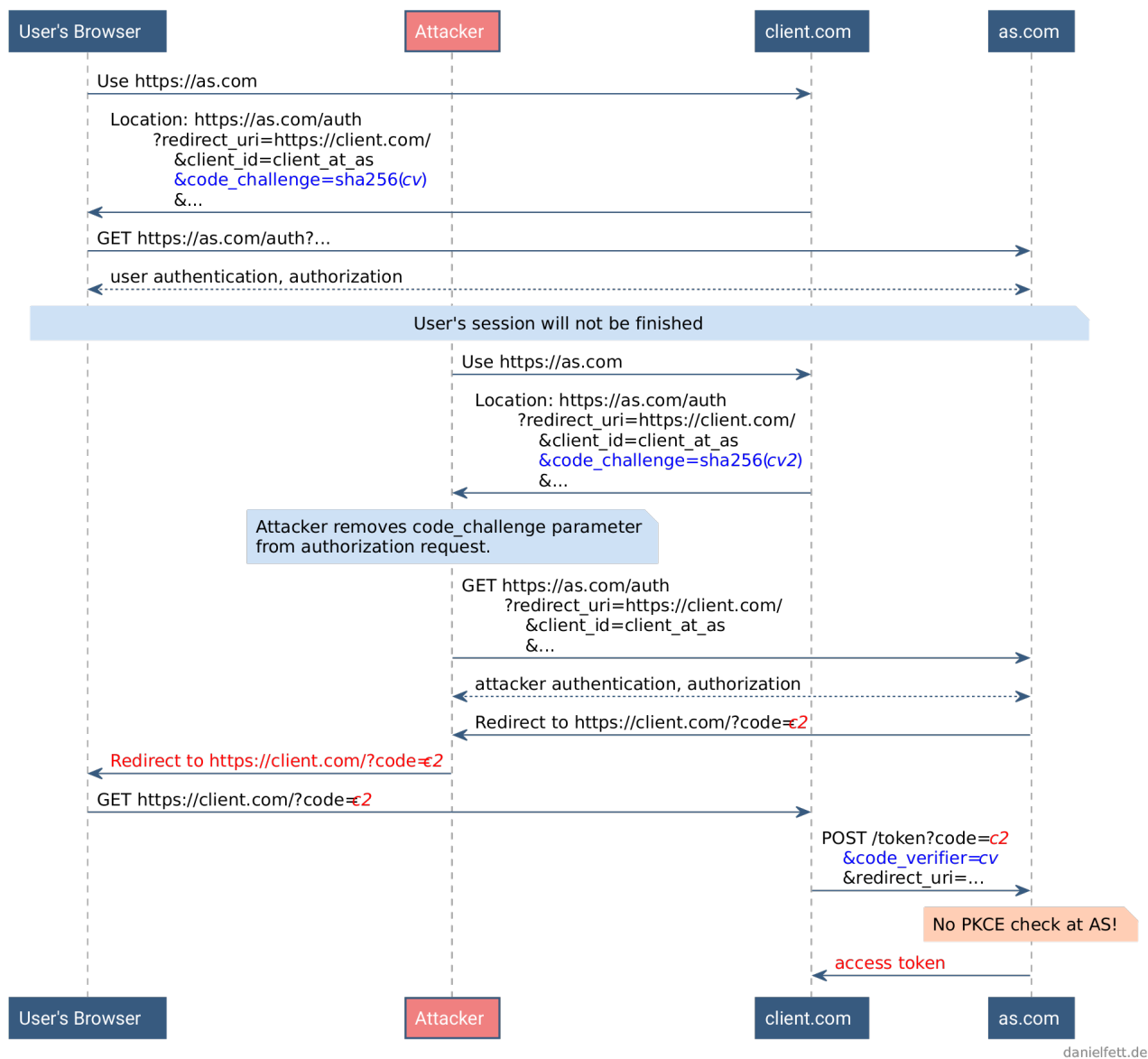


Figure 3.1: PKCE Downgrade. Source: Daniel Fett [9]

3.3 Language structure

Each part of the language is an Object, which is then written as a JSON Object in the language, that contains a series of name/value couples. The hierarchical structure of the language can be seen in Figure 3.2. An example of a Test Suite Object containing other Objects can be seen in Figure 3.3. In this section all the Objects composing the language will be discussed.

3.3.1 Test Suite

The Test Suite is the main Object which contains all the other one, it has these tags:

- **name**, the name of the test suite
- **description**, the description of the test suite
- **tests**, which is a list containing the tests to be executed

To implement the PKCE test example above, the Test Suite can be defined in this way:

```
1 {  
2   "test suite": {  
3     "name": "OAuth active tests",  
4     "description": "A test suite containing a OAuth test"  
5   },  
6   "tests": [  
7     {  
8       // A test  
9     }  
10  ]  
11 }
```

Listing 3.1: Test Suite definition

As shown in Listing 3.1, the Test Suite Object is a JSON object having the tag **test suite**, with name and description, and a tag **tests** which will contain a list of Test Objects.

3.3.2 Test

Following the hierarchical order, the Test object is the one that actually defines a test. As said earlier, a test is contained in a Test Suite, and contains various tags to be defined:

- **name**
- **description**
- **type**, it can be *active* or *passive*
- **sessions**, which is a list of the sessions which are needed in this test
- **result**, (only for actives) it defines the conditions over which the test is considered passed or not
- **operations**, a list of Operation objects which will be executed

A test can be defined either as active or passive depending on the type of actions it has to do on the intercepted messages. If a test does not need to manipulate the flow or the content of the messages is considered passive, otherwise it is considered active. The list of Operation Objects contained in a Test is executed iteratively one after the other, so, only after an Operation is completed, the next one starts.

Following the PKCE example introduced in Section 3.2, an active Test is defined:

```
1 {  
2   "test": {  
3     "name": "PKCE Downgrade",  
4     "description": "Tries to remove code_challenge parameter",
```

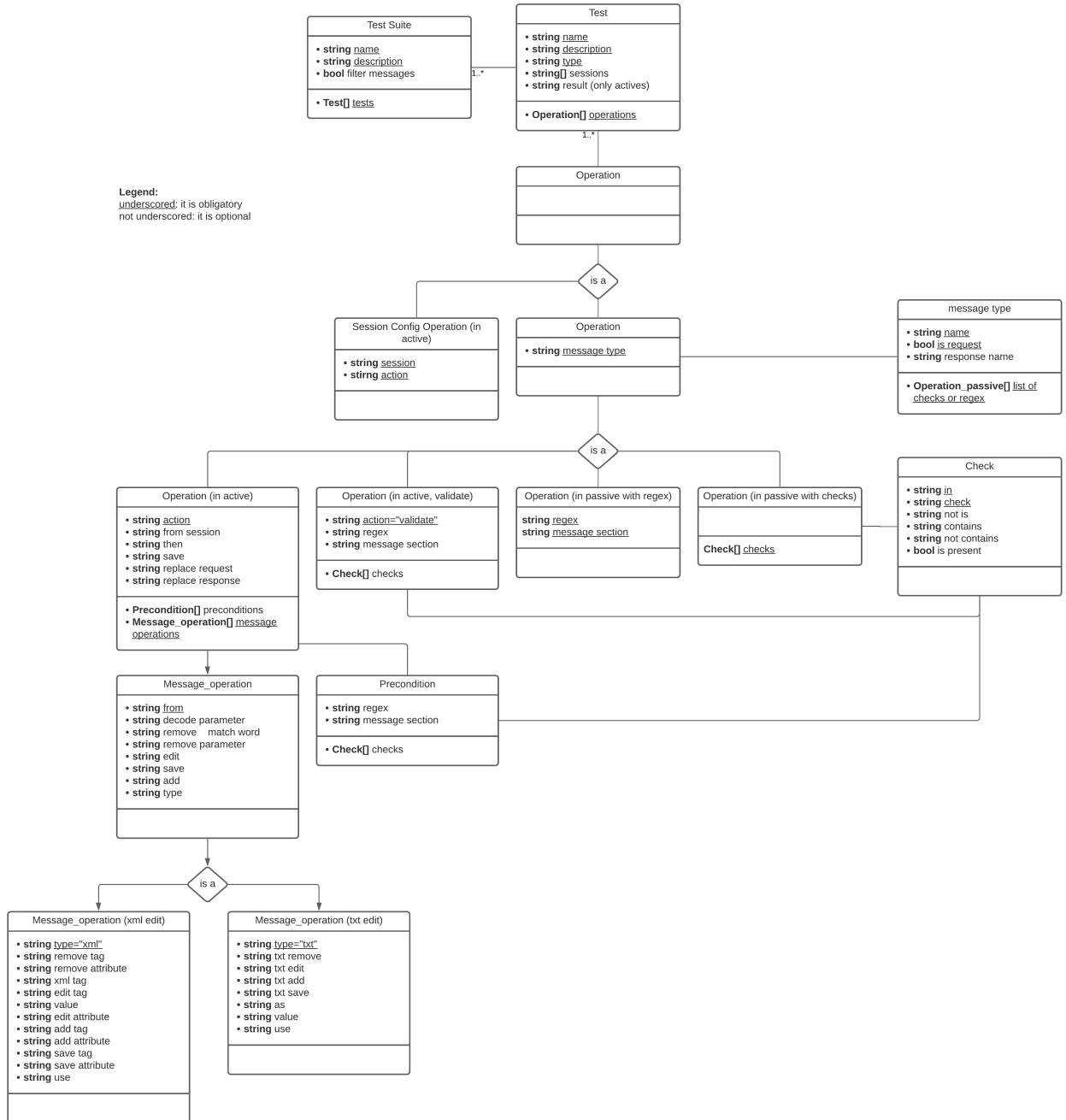


Figure 3.2: Language structure

```

5      "type": "active",
6      "sessions": [
7          "s1"
8      ],
9      "operations": [
10         // list of Operation Objects
11     ],
12     "result": "incorrect flow s1"
13 }
14 }

```

Listing 3.2: Active test definition

As can be seen in Listing 3.2, the test is specified, writing its type, which is active, as it has to edit a message removing a parameter from the url. Moreover, a session and a list of Operation to be executed are added. The result of the test is also specified.

3.3.3 Operation

The Operation Object defines the actions that a test has to do. As shown in the language structure in Figure 3.2, an operation could be either a **standard operation** or a **session config operation**, the latter is used to manage the sessions for the active tests (i.e. start, stop, pause them). Depending on the type of test which an Operation is defined into, the standard Operation becomes active or passive. In both cases, an operation has to contain the **message type** which defines the type of message to be intercepted in that particular operation (more info in the dedicated paragraph).

A **passive** operation has as objective to verify the presence (or absence) of some text or parameters in the intercepted message, to do this, it should contain one of the following options:

- A list of **Check** Objects, which are then executed to check the presence (or absence) of some text or parameter
- A **regex** inspection, which executes an inspection considering the intercepted message as plain text and executing a regex over it, if the regex has a match, the operation is considered passed, otherwise failed. Note that when a regex is used, it has to be specified also the **message section** over which to execute it (body, head, url)

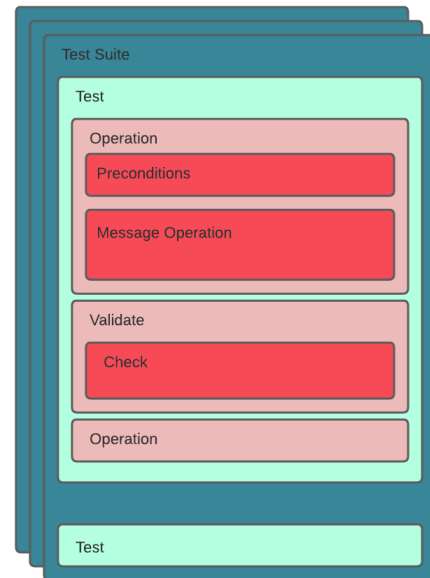


Figure 3.3: language structure

If the Test Object where the operations are defined is an **active** test, also, if the intercepted messages need to be manipulated, an active Operation has to be defined. It is composed by:

- **action**, the action it has to do (*intercept*, *validate*)
- **from session**, from which session to expect the message to be intercepted
- **then**, the action to do after the receiving and manipulation of the message (forward or drop)
- **replace request (or response)**, specify a previously saved message in order to replace it to the intercepted one
- **preconditions**, a list of Precondition objects

- **message operations**, a list of Message Operation objects, which will do the actual manipulation of the intercepted message

If the action is set to *validate* the operation is still active, but becomes like a passive operation, because its objective is just to verify that some messages meets some requirements. It will contain a regex or a list of checks to be done. The Validate Operation Object is part of the oracle, that is the component that decides whether the tests should be considered passed or not, more details will be discussed in the apposite section.

In the PKCE example seen in Section 3.2, an Operation Object for an active test is specified. This Operation has to intercept the *authorization request* message from session *s1*, it has to check some Preconditions over it and then do some Message Operations. At the end the message is forwarded to its original destination.

```

1 {
2   "action": "intercept",
3   "from session": "s1",
4   "then": "forward",
5   "message type": "authorization request",
6   "preconditions": [
7     // Precondition list
8   ],
9   "message operations": [
10    // Message Operation list
11  ]
12 }
```

Listing 3.3: Operation definition

3.3.4 Message section

The message section specifies in what part of the message execute the given action. According to [10] the parts that compose an HTTP message are:

- **url**, a start line describing the message
- **head**, a block of headers containing attributes
- **body**, an optional body containing data

3.3.5 Check Object

The Check Object is used in Operations Objects and in other Objects to verify that, in a message, some circumstances are satisfied. For instance, it can be used to check that a specific parameter in the url of a message should be equal to a specific string. The Check Object is defined by:

- **in**, where to search the given parameter. The possible value is *head*, *body* or *url*
- **check param**, specifies the parameter name to be searched

Also the actual checks on the parameter value has to be defined: (if none of these are defined, the Check will only check if the given parameter is present or not in the given section)

- **is**, checks that the parameter value is exactly what is passed to this tag
- **not is**, checks that the parameter value is not what is passed to this tag
- **contains**, checks that the parameter value contains what is passed to this tag
- **not contains**, checks that the parameter value does not contain what is passed to this tag
- **is present**, used to explicitly tell to just check the presence of the parameter, it accepts true or false, depending on if, respectively, the presence or the absence of the parameter has to be checked.

Note that by how the logic of the language has been thought, to consider a test passed, all the check Objects has to be evaluated to true. For instance, if a parameter should not be present in a given message, the check has to verify that the parameter is **not** present. If a Check is evaluated to false, the test will fail.

3.3.6 Precondition and Validate Objects

Check Objects in active tests will not work as in passive tests, there is another way of using them: using them in a Precondition, which is basically a list of the Check Objects, or by using the validate option in an Operation Object, which will make possible to use Checks to validate a given message. This difference of Check Objects between active and passive tests has been done to allow the Oracle to work, differentiating between a precondition and a validation. This way Validate Objects can be used to define the oracle, and Precondition Objects can be used to impose the criteria to allow the execution of the test.

Precondition Object Preconditions are used in an Operation of an active test to check that the intercepted message is compliant to specific criteria fixed by the tester, before the execution of the Message Operations. If the Check objects in the precondition are evaluated to false, the test is considered unsupported, not failed. More precisely the preconditions will be a list of Check Objects. This can be useful in case that it is not known if a given test is compatible with a given web service. For example if the service does not use some type of protocol, this can be assured this by the use of preconditions, checking for the presence of common parameters of the protocol. Preconditions can be also regex.

Validate Object In an Operation Object the only way to use Check objects is by setting the **action** tag to *validate*, this will transform the Operation in a Validate Operation. This Validate Operation will be used by the Oracle to decide whether the test should be considered passed or not. Validate Operation has to be used when a given part of a message should be in a specific way, if it is not, the result of the test will be considered failed. In order to do this the Validate Operation with a list of checks or a regex (exactly like in passive tests) can be used.

3.3.7 Save

A message or a string can be saved by the use of the tag **save**, this can be used both in an Operation, to save an entire message, or in a Message Operation to save the value of a found parameter. So a variable can be a message-type variable, containing an entire saved message, or it can be a string-type, containing a string. There are two ways of using the value of a variable which depends on its type:

- Using a **message-type** variable: it can be used in an Operation with the tag **action** set to intercept. There is the possibility of using **replace request** (or **replace response**) tag giving the name of the variable. This will replace the intercepted message's request (or response) with the message saved in the variable.
- Using a **string-type** variable: can be used in Message Operations, where a parameter has to be edited, writing **use** tag specifying the name of the variable to use. This will use the value in that variable in the way specified by the other tags (i.e. tags **edit** or **add**)

3.3.8 Message Operation

The Message Operation is the Object that actually does the manipulations on the intercepted messages. It is composed by these tags:

- **from**, the message section to work on
- **decode parameter** (optional) it indicates which parameter's value or string to be decoded before it can be processed
- **encodings** (optional) the list of encodings to be applied to the parameter or text to be decoded. The supported encodings are base64, deflate, url

- **remove match word** (optional), it accepts a regex, everything matched by that regex in the specified message section is removed
- **remove parameter**, it accepts a parameter name, it removes both the name and the value of that parameter from the specified section
- **edit**, it accepts a parameter name, and edits its value with the new value specified with **in** tag
- **save**, (optional) it accepts a parameter name, the value of that parameter is saved in a variable, it is necessary to specify the name of the variable using **as** tag
- **add**, (optional) it accepts a parameter name, it appends to the parameter value the string passed with **this** tag
- **type** (optional) specify how the decoded parameter should be interpreted (txt or xml)

In a message operation there is the possibility to specify a parameter or some text to be decoded before manipulation, to do that, specify with **decode parameter** the parameter to be decoded and with **encodings** the encodings necessary to decode the parameter. The order of definition of the **encodings** will be followed during decoding. The parameter (or text) decoded, at the end of the Message operation will be encoded again automatically before forwarding it. The decoded parameter can be manipulated by means of the **type** tag, there is the possibility to interpreter the decoded parameter by two means:

Type txt Associated to this interpretation, it is possible to use a list of actions over the plain text:

- txt remove: removes the matched string from the decoded parameter
- txt edit: edits the matched string with a custom string (specified with the **value** tag)
- txt add: after the matched string adds a string specified with the **value** tag
- txt save: saves the matched string in a variable with name specified in the **as** tag

All the previous tags accept a regex, and anything that regex matches, will be edited or added or saved based on the specified tag.

Type xml Another possibility is to interpret the decoded text as xml, to do this, the type tag has to be set to *xml*. This way the various possible operations to be done on the decoded xml are:

- **remove tag**, removes the specified tag
- **remove attribute**, removes the specified attribute associated with the xml tag specified using the **xml tag**
- **edit tag**, edits the specified tag with the value contained in **value** tag
- **edit attribute**, edits the specified attribute associated with the xml tag specified using the **xml tag**
- **add tag**, adds the specified tag, having the value specified with tag **value**, and also the name of the parent xml node to add the new node to, has to be specified using the **xml tag**
- **add attribute**, adds the specified attribute associated with the xml tag specified using the **xml tag**
- **save tag**, saves the specified tag value
- **save attribute**, saves the specified attribute value associated with the xml tag specified using the **xml tag**

In the PKCE example found in Section 3.2, a simple Message Operation is present, which has to remove the parameter `code_challenge` from the url of the message, so the resulting Message Operation will be:

```
1 {  
2   "from": "url",  
3   "remove parameter": "code_challenge"  
4 }
```

Listing 3.4: Message Operation definition

Note for body section in message operations If the *body* section is chosen, the meaning of the following tags becomes different:

- **remove parameter** will work like **remove match word**, in a way that the value of the tag is treated as a regex which will be matching against the entire body section, having all the matches removed from it
- **edit** is treated as a regex, substituting everything that matches that regex with the text specified by the **in** tag
- **save** is treated as a regex, saving what will be matched by the regex, the name of the variable in which the value will be saved is specified with the **as** tag
- **add** is associated with a regex, it will add at the end of the matched text the value specified by **this** tag

This changes also for the **decode parameter** tag, in a way that if the message section is 'body' the **decode param** will accept a regex, and everything matched by that regex will be considered to be decoded.

An example of a regex to match a parameter in the body could be `(?<=parameter_name=)[^\n&]*` that will search for the text "parameter_name", taking everything after the "=" until end of line or "&" or whitespace is found.

This difference of the tag meaning is due to the difficulty of identifying parameters in the body section in contrast to the head section. In fact, while the head section is based on the HTTP standard, having all parameters defined in a clear and well-defined way like "Name: value" the body section could contain any type of content. To manage this variety of contents the decision of using regex instead parameter names for the body section has been chosen.

3.3.9 Message Type definition

The message type definition is needed in order to define some types of message that will be later used in the language to intercept them. The message type definition is not actually part of the language, but it is stored in a file in the Burp folder. The definition of the type of messages uses the same Objects as the language. A Message Type Object is defined using these tags:

- **name**, the name that will be used in the language to refer to this message type
- **is request**, if set to true if the searched message is a request, false otherwise
- **response name**, if the searched message is a request message, this tag can be used to associate a name to the response of that message. This is useful when only the request message can be identified, making possible to intercept the correlated response message.
- **checks**, a list of Check objects used to identify the message. If evaluated to true, the message is considered found

Following the example from Section 3.2 the definition of the *authorization request* message used in the test, and can be found in the Listing 3.5.

```

1 {
2   "message_types": [
3     {
4       "name": "authorization request",
5       "is request": true,
6       "response name": "authorization response",
7       "checks": [
8         {
9           "in": "url",
10          "check param": "response_type",
11          "is present": "true"
12        }
13      ]
14    },
15  ]
16 }

```

Listing 3.5: Message Types definition

When used, this message type will search for the parameter `response_type` in the url of every request message. When a match is found, the Operation in which this Message Type is used will be executed.

3.4 The oracle

The ensemble of all parts of the language that decide the result of the tests is called Oracle, which decides whether a test should be considered passed or failed (or not applicable). I decided to build the oracle in a way that it can be almost fully customized by the user. The oracle is based on three main components:

- Evaluation of the complete (or incomplete) execution of the session track
- Evaluation of the Precondition objects
- Evaluation of the Validate objects

If all the above conditions are met, the test is considered passed, otherwise it is considered failed (or not applicable). The oracle can be built for example by using Validate objects to verify that some intercepted messages satisfy some conditions like having a particular parameter or string in them.

To build the oracle for the PKCE example mentioned in Section 3.2, both the result of the test and the precondition has been used, specifically, the precondition part is shown in Listing 3.6

```

1 "preconditions": [
2   {
3     "in": "url",
4     "check param": "code_challenge",
5     "is present": true
6   }
7 ],

```

Listing 3.6: Precondition definition

This precondition is used to consider the test “not applicable” if the parameter `code_challenge` is not found in the authorization request message. This means that is not possible to execute the given test over the actual intercepted message if the preconditions are not satisfied.

The **result** tag of the Test in the PKCE example in Section 3.2 is set to:

```

1 "result": "incorrect flow s1"

```

This means that the oracle will consider the test passed, if and only if the execution of the session track of the session named `s1` will be incorrect. The execution is considered incorrect when the execution of the session track fails, this can be for instance, because of an unexpected page is displayed, or an

element that should be clicked is not present in the page. In the case of the PKCE example in Section 3.2 the test is passed if the removal of the parameter `code_challenge` is not admitted by the AS, usually, if this is the case, an error page will be displayed.

3.5 Sessions

A session is a browser executing a session track, a session track is a list of user actions that the browser will simulate automatically during execution of the Tests. There is the possibility of defining and using more than one session, in a way that e.g., reply tests can be executed. Different sessions can have different session tracks.

As said in Section 3.3.3, a **from session** tag can be specified in the Operation, this will specify in which session to search the message. To define the session track the idea used in [11, 8], has been used, adding some options like **wait** and **clear cookies** functionalities. The syntax of the session track is based on the plain text export of Katalon Recorder [1]. An example of a session track that does the login at Unitn website is this:

```
1  open | https://www.google.com/ |
2  click | id=L2AGLb |
3  click | link=Accedi |
4  click | id=identifierId |
5  type | id=identifierId | matteo.bitussi@studenti.unitn.it
6  click | id=identifierNext |
7  click | id=clid |
8  type | id=clid | matteo.bitussi@unitn.it
9  click | id=inputPassword |
10 type | id=inputPassword | password
11 click | id=btnAccedi |
12 click | link=Gmail |
```

Listing 3.7: Session track Unitn login

This session track will do the login on the Unitn website using some credentials and password. The supported actions are:

- **open** | **url** |, to open an url
- **click** | **id=**, **link=**, **xpath=** |, to click on a http object with the given id, link or xpath
- **type** | **id=** | **text**, to write on a given http element the given text
- **wait** | **milliseconds**, to make the execution of the session wait for a given time
- **clear cookies** |, to make the browser of the session clear all the cookies in it

4 Implementation

In this chapter the implementation of the language and the tool, the problems faced, and the solutions adopted will be discussed

4.1 General overview of the tool

The components of the final tool can be seen in Figure 4.1. Burp Suite is composed by its proxy and related APIs, the tool will get all the messages from the proxy through the API, it will process them, and return them to the proxy. This way all the messages will pass through the tool, and will make possible to check and edit them. Every browser, one for each session, will be using a dedicated proxy, which will act like a Man-In-The-Middle attack from the browser to the server, establishing a secure connection only on the last part of the communication to the server, making possible to see plain HTTP communications on the browser side. Each browser will be supplied with the user actions which will be taken from the session tracks specified beforehand. It is also possible to do manual user actions on the browser, in case (for example) a captcha has to be resolved. Also, the session actions taken from the tests defined by the language can be supplied to the browser (for example to pause or stop it). The most important component of the tool is the Test Suite defined with the language, which is supplied to the plugin, and executed in ensemble with the sessions, giving eventually a result.

4.2 The tool

To implement the tool, I have decided to start from a work done by my colleague Wendy Barreto in her bachelor thesis [2], which realized a similar tool for OIDC and OAuth SSO protocols, this was a good base to start with my implementation. The interface of [2] has been taken and adapted to fit the needs of this work. The tool code is written in Java, I used the Burp's interface classes to interact with it. The standard usage of Burp is based on the execution of a browser which connects to the Burp's proxy, in a way that all the packets can be intercepted, viewed or edited and forwarded or dropped from the Burp interface. The tester would do some actions on the browser and watch the message flow in Burp and then check them or edit them. With the tool the idea is the same, but the operation done on the browser and the checks or edits on the messages are made automatically, in a way that the tester does not have to do them by itself.

4.2.1 User Interface of the tool

In Figure 4.2 the interface of the tool is shown, starting from top left, there is the session track input space, where it can be specified a different track for each session. Following on the top right, a series of buttons that allow various configurations can be found:

- **Use Chrome** or **Use Firefox** the browser to be used can be selected
- **Select driver** the driver used to automate the actions on the browser can be selected
- **Record** the record button can be used to record the flowing messages
- **Load messages** the load messages button can be used to load the previously saved messages to be tested offline
- **Offline mode** the offline mode button to test the loaded messages instead of the live ones
- **Execute track** used to execute the session track without executing the tests, useful when the unaltered messages has to be saved.
- **Test track** used to test the session track without saving or doing any test.

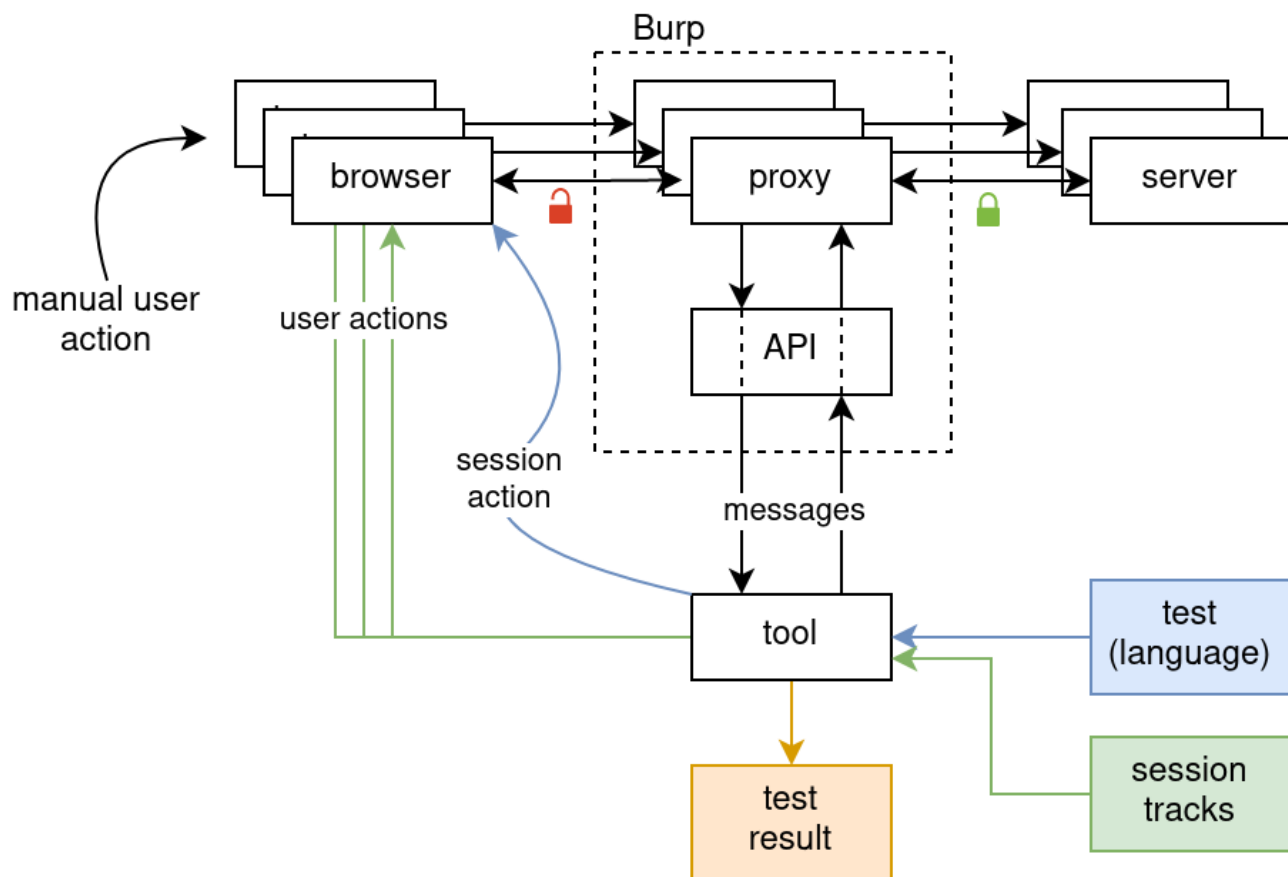


Figure 4.1: General schema

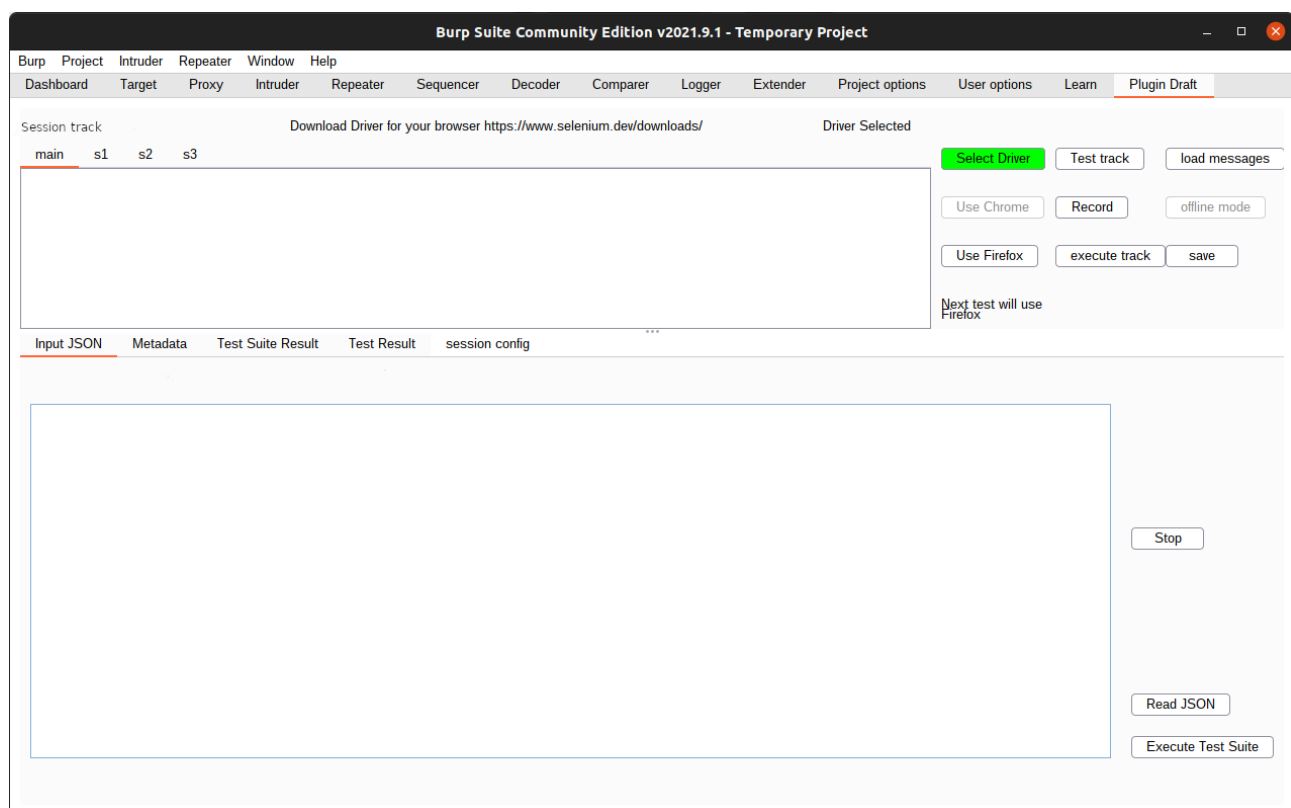


Figure 4.2: Tool interface

In the bottom part multiple tabs can be found:

- **Input JSON** tab is used to load the tests written in the language into the tool, and with the use of two buttons the language can be parsed, and the tests can be executed.
- **Test suite result** is the tab containing all the results of the executed tests
- **Test result** is the tab used to see the specific test result, with all the intercepted messages related to it
- **Session config** is used to configure the ports of the sessions that will be used in the tests.

In the bottom right part, when the **Input JSON** tab is selected, three buttons are available:

- **Stop** used to stop the current execution
- **Read JSON** used to read the written Tests
- **Execute Test Suite** used to execute the written tests

4.2.2 Session managing

The sessions are managed independently, each session is basically a browser that is launched when a session is started. Each session can follow a different session track defined in the apposite tabs. Every session is run in a separated thread to make parallelism between every session possible. By the use of specific commands in the language, is possible to do some actions on each session, like stop it, pause it, or clear its cookies. Each browser uses a different proxy port, so that it is possible to know from which session the messages come from, and so, being able to specific sessions in the tests.

4.2.3 Test execution

The test execution differs from passive to active, as passive tests do not need the edit of the messages, the execution of the session track is done once, the messages are saved and the tests are executed on the saved messages. I have also added the possibility of exporting the saved messages to a file, in a way that they can be imported in the tool and tested again. On the other hand, active tests needs to edit the messages, so the execution of the track has to be repeated for each test.

4.2.4 Decoding & encoding of parameters

As said in Chapter 3, the decoding and encoding of parameters is possible. To do that, a list of encodings to be done on the parameter has to be provided, e.g., url, base64, deflate. Once the specified message is intercepted, the parameter is taken and decoded following the order of the provided encodings list. To do that, part of the code of SAML Raider [16] has been used. SAML Raider is a Burp's plugin used to manage SAML certificates. The part of the code which deals with encoding and decoding of the parameters has been taken and edited to fit the tool.

4.2.5 SAML certificate managing

In SAML Requests and responses there is sometime the need to remove or edit the certificate associated to that request or response, so, to speed up the process a specific tag in the language has been added to remove or edit the certificate signature. There is still the possibility of doing the same removal by editing the SAML request or response with a regex, but with the use of the tag this becomes more convenient. In order to do this, a part of the code of SAML Raider [16] has been used, and edited to fit the needs of the tool.

4.3 Problems and limitations encountered

During the implementation and the testing of the tool multiple problems have been encountered, the majority of them have been solved, but some are still present. The most relevant ones will be discussed next:

4.3.1 Automation problems

One of the biggest limitations that the tool has, is the session track actions automations, it often happens that some captcha is encountered during execution, making impossible to proceed. Moreover, the track execution is limited, there is only a possible flow of actions (the one defined) and there is not the possibility of inserting “if then else” constructs that could help to differentiate the actions based on the actual page or popup. For example, it could happen that a “limited time offer” popup could appear in a website only in a particular time, the execution of the session track could be compromised by that, making impossible to distinguish whether the test failed because of the tested vulnerability or the actual popup. Another problem in the automation part of the tool is that is sometimes limited, because the session track has to be defined over a specific website, doing a set of action that is directly correlated to the website. Whenever the website is changed somehow, for example the IDs or the position of some button to be clicked change, the execution will fail, because the track could not continue. This is still not resolved, as no methods to make the track more dynamic has been found yet. This also means that every different web service which has to be tested will need a different session track to be defined. Making it a bit time-consuming to do.

This Automation problems mainly occur when the service to be tested is not under the tester control, otherwise, captcha and other popup could be easily deactivated for testing purposes. This means that this problem is happening only on specific use case scenarios, and so, it is not a big deal.

4.3.2 Oracle is sometimes ambiguous

There still is a problem with the Oracle, where sometimes false positives or negatives arise if the execution of the tool is interrupted for any reason. This is a problem because the interruption of the execution is a term of valuation for the Oracle, this means that the oracle will give a result also based on the correctness or incorrectness of the execution of the session track. This makes impossible to distinguish if the session track has failed because on an error on the definition on it, or because of an expected reason (like after a message modification).

4.3.3 Interface and user feedbacks

The interface of the tool is a bit raw, it is not very user-friendly, the user experience could be improved. During my work I did not focus my attention to these topics, but they are very important as the tool is not so easy to use. Also, the feedbacks of the errors encountered by the tool such as execution errors or others are not all shown to the user, this surely has to be fixed, making more clear to the end user what is going wrong.

5 Use cases

In this chapter some examples of use cases in which the work of this thesis has been used will be shown.

5.1 SAML Use-Case

During the last stages of the development and testing of the tool, a strict collaboration between my colleague Sofia Zanrosso and me has started. Her objective was to create a SAML Test Suite in order to facilitate automatic pentesting over SAML[17]. The tool defined in this thesis was compared with other ones and was used to define and execute the tests. During the progress of both our works, a lot of feedbacks and bugs has been reported to me, speeding up the testing phase of the tool. At the same time, we found that my tool was in some respects better compared to the other alternatives. For example:

- Other plugins were giving false positives on some tests
- "the previously employed transition times between tools have been greatly reduced"
- "making it possible to analyze almost completely the vulnerabilities of the tested subjects"

5.2 OAuth & OIDC Use-Case

The OAuth and OIDC tests defined and used in [11, 2] has been re-defined in the language and has been used, during development, to evaluate the correctness of the tool with respect to the already existing ones. The tests written in SBTL language and the other had the same results. Due to the new available functionalities introduced with SBTL, new tests has been specified, especially active ones. At the end of the development, a Test Suite written in SBTL for OAuth and OIDC has been defined.

6 Related works

This chapter will describe software and tools that are related to the work in this thesis.

6.1 Micro-Id-Gym

Micro ID Gym (MIG) “aims to assist system administrators and testers in the deployment and pen-testing of IdM protocol instances” - [3], inside this tool, two pentesting tools can be found:

- MIG - OAuth/OIDC [11]
- MIG - SAML SSO [8]

They are both plugins for Burp, which have as objective to test the two different protocols. These plugins execute a series of actions on a browser, checks the messages in the background, and then, they provide a result. These two plugins are related to mine, but they are specifically created and defined to test SSO protocols only, and the tests that they used are fixed and cannot be easily edited, if a new test has to be implemented, the plugin has to be recompiled.

6.2 SSO Testing language and Plugin

The preceding two tools of MIG found in Section 6.1 have been improved by the work done by my colleague Wendy Barreto [2] in her bachelor thesis at Università di Trento to test OAuth and OIDC SSO protocols with a custom test definition pattern. Her work aimed to fix the problem of hard-coded tests in the plugins for SSO protocols testing. The previous MIG plugin had been improved by removing the staticity of the test, adding the possibility to customly define all the tests with the use of a JSON language. The available test actions worked well, but there were some limitations on the possible actions, especially in the active tests. For example:

- Limited oracle for the verification of active tests, having just the verification of the correct execution of the operation and a check for the string “error” on the last page of the browser
- The filtering of the message to check or edit for static tests is limited, only *Authorization grant message*, *Response messages*, *Request messages* and *All messages* are available
- Only regex are supported to search what is needed in a message
- Impossibility to work over encoded parameters
- Impossibility of doing multiple operations on a single message
- Impossibility of saving a parameter and using it somewhere else
- Impossibility of using multiple sessions in a test

Some of these limitations were stated as future works in [2]. A more complete list of differences between the two languages and tools can be found in Attachment B.

Previously in this thesis it has been said that part of this work has been taken as a base to start with the implementation of the tool. The idea of a language that could be used for any type of test over HTTP was born when I used her plugin, which was limited to OAuth and OIDC tests. I wanted to enlarge the possible tests to be defined without a restriction on a specific protocol. One of the things that has been used, is the interface of the plugin, that has been modified, adding buttons and tabs to deal with multiple session tracks and other added functionalities. Also, the automation of the session track was taken and edited, this part was already used in [11, 8].

7 Conclusions and Future Works

The objective of this work, was to design and implement an automated tool that could execute security tests specified with a language, making possible for the tester to use an already defined test suite, saving time and being sure that all the tests for a specific protocol are done. This objective can be considered reached, the new tool and language is working, as can be seen in [17], in this case, all the necessary tests have been specified in the language, and have been executed, having the expected result.

Still, it is not perfect in terms of stability and affidability, as a deep testing phase should be accomplished, but overall, the tool is functioning well.

For future works, the problems and the limitations seen in Section 4.3 should be solved. The message filtering part of the language could be extended by the use of an AI, making it possible to define a more abstract filter that does not solely rely on the search of parameters or exact strings. Another thing is that the test definition language could be extended even for other uses, like for low-level networking protocols such as routing protocols, electronic trading protocols, IP protocols, and so on. Making it possible to search for known vulnerabilities in real-time or just by analyzing saved packets files. For example, a plugin for Wireshark software could be developed.

Bibliography

- [1] Katalon recorder sample plugin. <https://github.com/katalon-studio/katalon-recorder-sample-plugin>. Last accessed 19/11/2021.
- [2] Wendy Barreto. Design and implementation of an attack pattern language for the automated pentesting of OAuth/OIDC deployments, 2019/2020. Bachelor thesis.
- [3] Andrea Bisegna, Roberto Carbone, Ivan Martini, Valentina Odorizzi, Giulio Pellizzari, and Silvio Ranise. Micro-id-gym: identity management workouts with container-based microservices. *Int. J. Inf. Secur. Cybercrime*, 8(1):45–50, 2019.
- [4] Tim Bray. The JavaScript Object Notation (JSON) Data Interchange Format. RFC 8259, December 2017.
- [5] Brian Campbell, Chuck Mortimore, and M Jones. Security Assertion Markup Language (SAML) 2.0 profile for OAuth 2.0 client authentication and authorization grants. *Internet Engineering Task Force (IETF)*, 2015. Page 2, Last accessed 29/11/2021.
- [6] Yulia Cherdantseva and Jeremy Hilton. A reference model of information assurance & security. In *2013 International Conference on Availability, Reliability and Security*, page 547. IEEE, 2013.
- [7] Milena Gabanelli e Simona Ravizza. Attacchi hacker, dati sanitari in pericolo: la lista segreta dei 35 ospedali colpiti. <https://www.corriere.it/dataroom-milena-gabanelli/attacchi-hacker-dati-sanitari-pericolo-lista-segreta-35-ospedali-colpiti/1abf2704-2079-11ec-924f-1ddd15bf71fa-va.shtml>, 2021. Last accessed 29/11/2021.
- [8] Stefano Facchini. Design and implementation of an automated tool for checking SAML SSO vulnerabilities and SPID compliance, 2019/2020. Bachelor thesis.
- [9] Daniel Fett. Downgrade attack on pkce. <https://danielfett.de/2020/05/16/pkce-vs-nonce-equivalent-or-not/>. Last visited 16/02/2022.
- [10] David Gourley, Brian Totty, Marjorie Sayer, Anshu Aggarwal, and Sailu Reddy. *HTTP: the definitive guide*. O'Reilly Media, Inc., 2002.
- [11] Claudio Grisenti. A pentesting tool for OAuth and OIDC deployments, 2019/2020. Bachelor thesis.
- [12] Dick Hardt et al. The OAuth 2.0 authorization framework. <https://datatracker.ietf.org/doc/html/rfc6749.html>, 2012. Page 1, Last accessed 29/11/2021.
- [13] Okta. Protecting mobile apps with PKCE. <https://www.oauth.com/oauth2-servers/pkce/>, 2015.
- [14] Il Fatto Quotidiano. Attacco hacker lazio, al poliambulatorio santa caterina cittadini rimandati indietro: “impossibile fissare visite o cambiare medico”. <https://www.ilfattoquotidiano.it/2021/08/03/attacco-hacker-lazio-al-poliambulatorio-santa-caterina-cittadini-rimandati-indietro-impossibile-fissare-visite-o-cambiare-medico/6282342/>, 2021. Last accessed 29/11/2021.

- [15] Nat Sakimura, John Bradley, and Naveen Agarwal. Proof Key for Code Exchange by OAuth Public Clients. RFC 7636, September 2015.
- [16] Compass Security. SAML raider plugin. <https://github.com/CompassSecurity/SAMLRaider>. Last accessed 23/11/2021.
- [17] Sofia Zanrosso. Development of saml testsuite to facilitate automatic pentesting, 2020/2021. Bachelor thesis.

Attachment A PKCE test Example complete

```
1 {
2   "test suite": {
3     "name": "OAuth Active tests",
4     "description": "A series of tests to test OAuth's well-known
↪ vulnerabilities",
5     "filter messages": true
6   },
7   "tests": [
8     {
9       "test": {
10         "name": "PKCE Downgrade",
11         "description": "Tries to remove code_challenge parameter",
12         "type": "active",
13         "sessions": [
14           "s1"
15         ],
16         "operations": [
17           {
18             "session": "s1",
19             "action": "start"
20           },
21           {
22             "action": "intercept",
23             "from session": "s1",
24             "then": "forward",
25             "message type": "authorization request",
26             "preconditions": [
27               {
28                 "in": "url",
29                 "check param": "code_challenge",
30                 "is present": true
31               }
32             ],
33             "message operations": [
34               {
35                 "from": "url",
36                 "remove parameter": "code_challenge"
37               }
38             ]
39           }
40         ],
41         "result": "incorrect flow s1"
42       }
43     }
44   ]
45 }
```


Attachment B Language comparison

Action	Old language	New language
Custom message filtering	Only on active tests	supported
Edit string	only by regex	supported with regex and check
Remove string	only by regex	supported with regex and check
Add string	not supported	supported
Check parameter	only with regex	with regex and check construct
Multiple operations in single message	not supported	supported
Saving and reusing of values and messages	not supported	supported
Multiple sessions in single test	not supported	supported
Custom oracle definition	not supported	by using regex and checks