A dark blue vertical bar on the left side of the slide. A blue arrow points to the right from the bar, containing the date.

30/03/2018

Creation of a module in AIL-framework

Discovery of AIL and creation of a
game related module

Several thin, curved lines in dark blue and light grey originate from the bottom left and curve upwards and to the right.

Kowalski Laurent and Kieffer Joris

M2 SSI METZ 2017/2018

SUMMARY

Introduction	2
Discovery of AIL-Framework	3
Introduction	3
Features	4
Creation of the module.....	7
Corpus of game related word/ expressions.....	7
The module and Its integration in the ail framework.....	9
Results.....	14
Conclusion.....	17

INTRODUCTION

During our last year of master's Degree in IT Security we had a course about Cyber Threat Intelligence with the professional speaker M. A. Dulaunoy.

We were asked to create a project about a subject or a software in IT Security. During this course we had the chance to discover different kind of software like MISP or AIL-framework. We also discovered other subjects like Malware classifier or how to recover data from a source.

We had an interest in the process of analyzing pastes to find data leaks or potential threats that can happen in the future.

So, we decided to work on AIL-framework. We were told that in these pastes they were a lot of useless data like Game conversations or Game related subjects. With this idea, we decided to create a module to find all these pastes containing Game related words and integer it into AIL-framework.

This report will be divided in 3 parts:

In the beginning we will talk about our discovery of AIL-framework, then we will talk about the module created and finally we will talk about the results of this module.



LOGO OF AIL-FRAMEWORK

DISCOVERY OF AIL-FRAMEWORK

INTRODUCTION

AIL is a modular framework to analyse potential information leaks from unstructured data sources like pastes from Pastebin or similar services or unstructured data streams. AIL framework is flexible and can be extended to support other functionalities to mine or process sensitive information.

<https://github.com/CIRCL/AIL-framework>

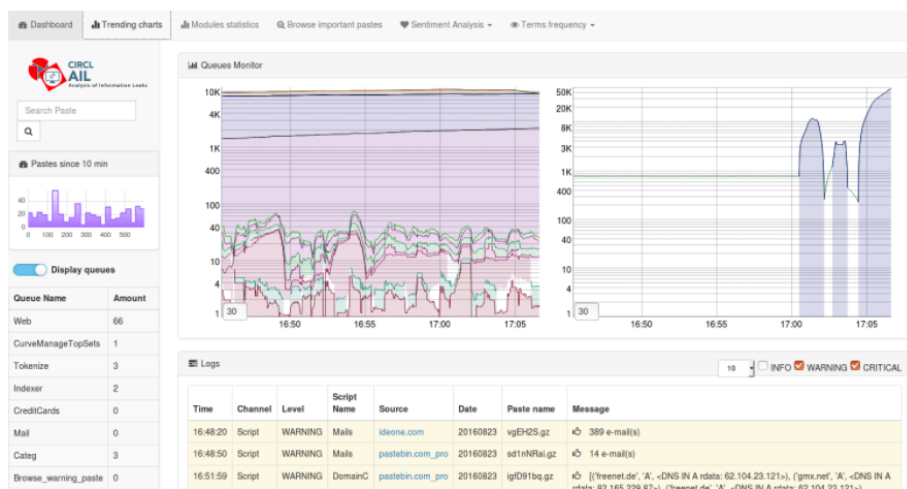
This framework was created by the CIRCL (Computer Incident Response Center Luxembourg) and you can use it freely and modify it.

AIL



AIL framework - Framework for Analysis of Information Leaks

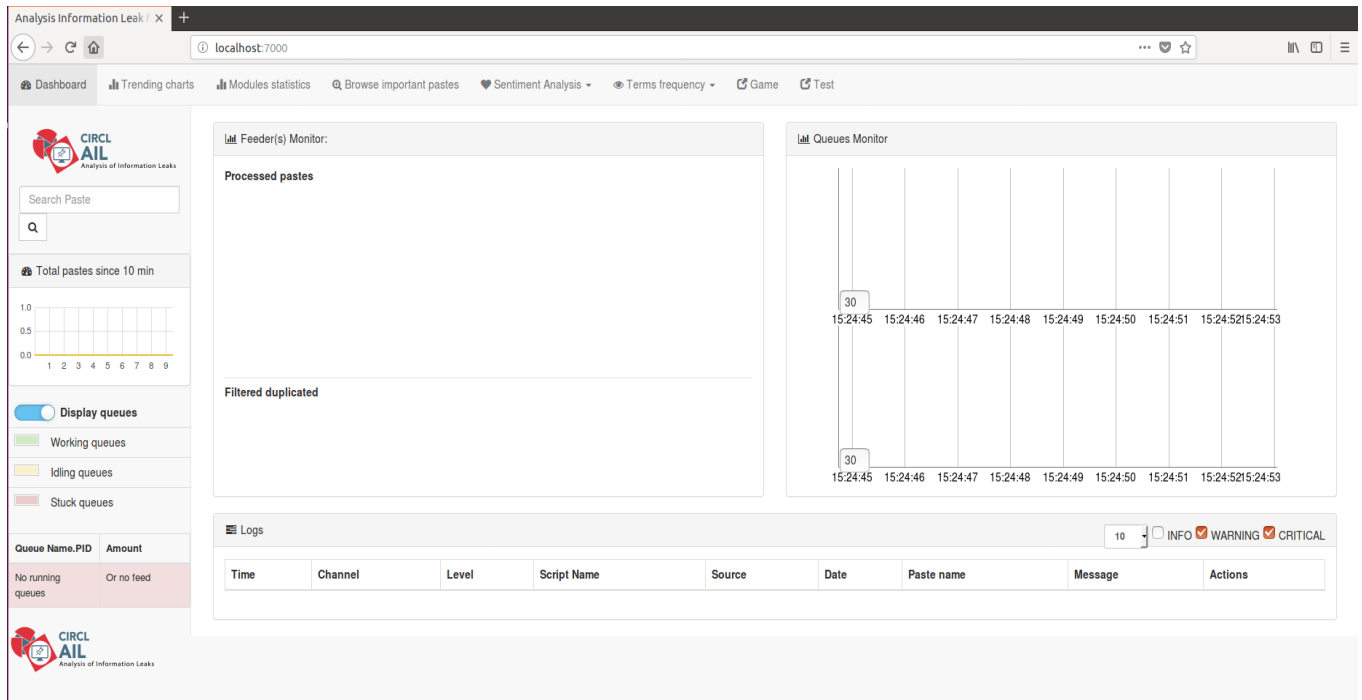
AIL is a modular framework to analyse potential information leaks from unstructured data sources like pastes from Pastebin or similar services or unstructured data streams. AIL framework is flexible and can be extended to support other functionalities to mine or process sensitive information.



FRAMEWORK

FEATURES

- AIL-framework uses modules to analyze different kind of data. You can for example search for credentials or email addresses or even phone numbers.
- It has a web interface allowing you to watch the process and the data found. You can even find graphics about the result of the modules.



WEB INTERFACE

- This web interface uses the local port 7000.
- On this interface you can see the pastes processed and what the framework found.
- There is a way to watch how the modules are working in 2 different python files: ModuleInformation.py and ModulesInformationV2.py. With these 2 files you can see if a module is working and start it or stop it as you want. It is simple to watch the modules.
- Alerting to MISP to share found leaks within a threat intelligence platform using MISP standard

Running Queues									
Action	Queue name	PID	#	S Time	R Time	Processed element	CPU %	Mem %	Avg CPU%
Idling Queues									
Action	Queue	PID	Idle Time	Last paste hash	Action	Queue	State	Queues not running	
					<S>	Attributes	No data		
					<S>	Categ	No data		
					<S>	Credential	No data		
					<S>	CreditCards	No data		
					<S>	Curve	No data		
					<S>	CurveManageTopSets	Not running by default		
					<S>	Cve	No data		
					<S>	DomClassifier	Not running by default		
					<S>	DumpValidDomain	Not running by default		
					<S>	Duplicates	No data		
					<S>	Game	No data		
Logs									
	Time	Module	PID	Info					

5

- The second interface is easier to start or stop a module.
- You can use different kind of sources for your analysis. You can use live feed from a source you had access. Or you can upload your own files using the import file. For this project we will use the import file with a dataset we got containing the pastes of December 2017.
- Every action or information found by the modules is logged into files in ALL-framework/logs
- You can create your own modules using the template they give you. And that's what we will do now.

- The list of words used for the detection and the false positives that can result from a bad list
- The module and it's integration in the framework

First, we needed to see if a list of this terms was already available online, so we found this page which contains a lot of game related terms:

[illegible]

We decided to put it at 5 occurrences of different game words because the goal is to find real data about games and not a false positive with 20 occurrences of the same word

We also thought which kind of users/gamers will use pastes to give information and after some first analysis we noticed that most of the files were configuration file for games and MMO related conversation, so we added a list of words from the world of MMORPG and for the configuration file they were already in the list.

When we are writing this report, the list is made of about 250+ words/expressions/games

Achievement
Across the map
Action-RPG
Add-on
Advergame
AFK
Aggro
Agilité
Aliasing
Anisotrope
Anti-aliasing
ARG
Artwork
Assassin
Beat them all
Bêta
Bonus
Boost
Bootleg
Boss
Bot
Bug
Bullet time
Burst
Camping
Carte
Casu
cel-shading
Cheat
Cheateur

EXAMPLE OF THE WORDS YOU CAN FIND FOR THIS MODULE

All these words can be found in the corpus.txt file in the github, this list can be modified to be more accurate or for another subject other than games.

THE MODULE AND ITS INTEGRATION IN THE AIL FRAMEWORK

We installed AIL framework on linux x64 virtual machine using the following tutorial:

Setting up AIL-Framework from source

```
1 git clone https://github.com/CIRCL/AIL-framework.git
2 cd AIL-framework
3 ./installing_deps.sh
4 cd var/www/
5 ./update_thirdparty.sh
```

TUTORIAL TO INSTALL AIL

The first step to create a module can be found on the howto.md file from the AIL-framework github page:

How to create a new module

If you want to add a new processing or analysis module in AIL, follow these simple steps:

1. Add your module name in `./bin/packages/modules.cfg` and subscribe to at least one module at minimum (Usually, `Redis_Global`).
2. Use `./bin/template.py` as a sample module and create a new file in `bin/` with the module name used in the `modules.cfg` configuration.

HOW TO CREATE A MODULE

So, we added these lines in the `modules.cfg` file:

```
[Game]
subscribe = Redis_Global
```

MODULES.CFG

So, our module will be called `Game` and will use `Redis_Global`. For now, the module will not send anything to other modules but the idea of using the `Game` module as a filter before other modules is an idea we thought about.

For the module file, we will use the template and just modify what we need to make it work:

```
1 #!/usr/bin/env python2
2 # -*-coding:UTF-8 -*
3 """
4     Template for new modules
5 """
6
7 import time
8 from pubsublogger import publisher
9
10 from Helper import Process
11
12
13 def do_something(message):
14     return None
15
16 if __name__ == '__main__':
17     # If you wish to use an other port of channel, do not forget to run a subscriber accordingly (see launch_logs.sh)
18     # Port of the redis instance used by pubsublogger
19     publisher.port = 6380
20     # Script is the default channel used for the modules.
21     publisher.channel = 'Script'
22
23     # Section name in bin/packages/modules.cfg
24     config_section = '<section name>'
25
26     # Setup the I/O queues
27     p = Process(config_section)
28
29     # Sent to the logging a description of the module
30     publisher.info("<description of the module>")
31
32     # Endless loop getting messages from the input queue
33     while True:
34         # Get one message from the input queue
35         message = p.get_from_set()
36         if message is None:
37             publisher.debug("{} queue is empty, waiting".format(config_section))
38             time.sleep(1)
39             continue
40
41         # Do something with the message from the queue
42         something_has_been_done = do_something(message)
43
44         # (Optional) Send that thing to the next queue
45         p.populate_set_out(something_has_been_done)
```

TEMPLATE

To make this module we will use regex expressions with every word in the corpus.txt file. And in order to stock the results we copy the file found and the number of occurrences in stock.txt

```

"""
Game module
"""

import time
import re
from pubsublogger import publisher
from packages import Paste
from Helper import Process
from pubsublogger import publisher

def search_game(message):
    #We recover the paste
    paste= Paste.Paste(message)
    content=paste.get_p_content()
    #We open the file with all game word and the stock for all paste found
    filetoopen=open("corpus.txt","r")
    filetowrite=open("stock.txt","a")
    count=0 #Number of different game word found in 1 file
    for line in filetoopen:
        line=line.strip() #Must do because it takes all the line and not just the word
        reg=re.compile(r'{}'.format(line.strip()))#We create the regex
        results=re.findall(reg,content)#We find the occurrences
        if(len(results)>0):#If the word is present in the paste ->+1
            count=count+1
        re.purge()
    if count>5:
        print results
        publisher.warning('{} contains Game related conversations+{} occurrences of a game related word {}'.format(paste.p_name,count))#warning for the logs
        filetowrite.write('{} contains Game related conversations+{} occurrences of a game related word {}'.format(paste.p_name,count))#For stock.txt
        to_print = 'GameConv;{};{};{};{} Terms related;{}'.format(paste.p_source, paste.p_date, paste.p_name, count, paste.p_path)#To see on the webinterface
        publisher.warning(to_print)
        filetoopen.close()
        filetowrite.close()

if __name__ == '__main__':
    # If you wish to use an other port of channel, do not forget to run a subscriber accordingly (see launch_logs.sh)
    # Port of the redis instance used by pubsublogger
    publisher.port = 6380
    # Script is the default channel used for the modules.
    publisher.channel = 'Script'

    # Section name in bin/packages/modules.cfg
    config_section = 'Game'

```

FINAL CODE OF GAME.PY

To use it now, it wasn't enough to just modify these 2 files. We needed to modify another file to make it work with AIL.

First, we added our module in both launch files for AIL-framework: LAUNCH.sh and launch_scripts.sh:

```

screen -S "Script_AIL" -X screen -t "SetForTermsFrequency" bash -c './SetForTermsFrequency.py; read x'
sleep 0.1
screen -S "Script_AIL" -X screen -t "Indexer" bash -c './Indexer.py; read x'
sleep 0.1
screen -S "Script_AIL" -X screen -t "Keys" bash -c './Keys.py; read x'
sleep 0.1
screen -S "Script_AIL" -X screen -t "Phone" bash -c './Phone.py; read x'
sleep 0.1
screen -S "Script_AIL" -X screen -t "Game" bash -c './Game.py; read x'
sleep 0.1
screen -S "Script_AIL" -X screen -t "Release" bash -c './Release.py; read x'
sleep 0.1

```

LAUNCH.SH

```

screen -S "Script" -X screen -t "Keys" bash -c './Keys.py; read x'
sleep 0.1
screen -S "Script" -X screen -t "Phone" bash -c './Phone.py; read x'
sleep 0.1
screen -S "Script" -X screen -t "Game" bash -c './Game.py; read x'
sleep 0.1

```

LAUNCH_SCRIPTS.SH

And finally, we needed to add our module into the doc to see it with the appropriate tools (ModuleInformation.py...)

```
Mixer
Global
PreProcessFeed
Duplicates
Indexer
Attributes
Lines
DomClassifier
Tokenize
Curve
RegexForTermsFrequency
SetForTermsFrequency|
CurveManageTopSets
Categ
CreditCards
Mail
Onion
DumpValidOnion
Web
WebStats
SQLInjectionDetection
ModuleStats
alertHandler
SentimentAnalysis
Release
Credential
Cve
Phone
Keys
Game
```

FICHER ALL_MODULES.TXT

Results: we can see our module in the tools

Running Queues							
Action	Queue name	PID	#	S Time	R Time	Processed element	CPU %
<K>	Web	17474	17	2018-03-30 20:32:34	0:01:25	1A1Pkyvp	0.00%
<K>	WebStats	17581	3	2018-03-30 20:31:45	0:02:14	fkh6pkjZ	0.00%

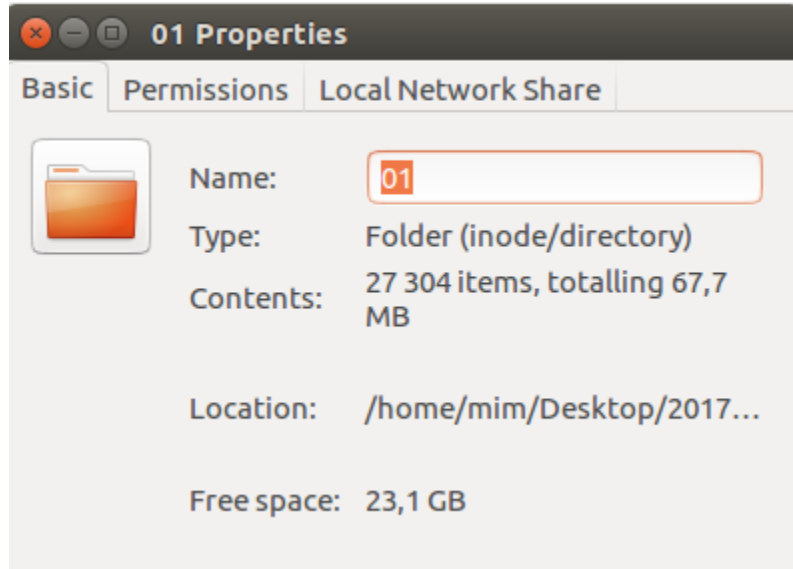
Idling Queues				Queues not running		
Action	Queue	PID	Idle Time	Last paste hash	Action	Queue
<K>	Attributes	17396	0:10:38	TCN95a4V	<S>	CurveManageTopSets
<K>	Categ	17430	0:10:37	LS2q07N3	<S>	Cve
<K>	Credential	17484	0:10:31	wYn24LgP	<S>	DomClassifier
<K>	CreditCards	17446	0:10:32	zaNU0Nmt	<S>	DumpValidOnion
<K>	Curve	17498	0:10:23	WkWcwqTS	<S>	Duplicates
<K>	Global	17367	0:10:39	abBhKgCz	<S>	Game
<K>	Indexer	17526	0:10:35	47x1b3fU	<S>	PreProcessFeed
<K>	Keys	17536	0:10:38	abBhKgCz	<S>	alertHandler
<K>	Lines	17408	0:10:32	706wHaV2		

		State	
		Not running by default	
		Stuck or idle, restarting disabled	
		Stuck or idle, restarting disabled	
		Not running by default	
		Stuck or idle, restarting disabled	
		Stuck or idle, restarting disabled	
		Not running by default	
		Stuck or idle, restarting disabled	

OUR MODULE WITH MODULESINFORMATIONV2.PY

RESULTS

To test our module, we had some pastes from December 2017. So, we will test this on the paste of 1 day



NUMBER OF FILES FOR 1 DAY WORTH OF PASTES

And now we will see how many pastes are related to games, and we will see some of these files.

After 1 hour, we decided to stop the analysis. AIL-framework analyzed around 3 000 pastes

Time	Channel	Level	Script Name	Source	Date	Paste name	Message	Actions
22:20:18	Script	WARNING	GameConv	Desktop	20171201	hjvxMComp.gz	Terms related	Q
22:20:26	Script	WARNING	GameConv	Desktop	20171201	sVUEzTY3.gz	Terms related	Q
22:20:29	Script	WARNING	GameConv	Desktop	20171201	71nc01bz.gz	Terms related	Q
22:20:29	Script	WARNING	GameConv	Desktop	20171201	k0ZR2aG5.gz	Terms related	Q
22:20:32	Script	WARNING	Credential	Desktop	20171201	BPM97dvz.gz	41 credentials found.	Q
22:20:32	Script	WARNING	GameConv	Desktop	20171201	EaduRKlw.gz	Terms related	Q
22:20:33	Script	WARNING	GameConv	Desktop	20171201	ZxNtzzJd.gz	Terms related	Q
22:20:35	Script	WARNING	GameConv	Desktop	20171201	P5KfI8ga.gz	Terms related	Q
22:20:44	Script	WARNING	Credential	Desktop	20171201	R9LddUli.gz	1372 credentials found.	Q
22:20:46	Script	WARNING	Credential	Desktop	20171201	v3JM4q2E.gz	1452 credentials found.	Q

RESULTS

Our module found around 892 pastes with game related words in it

Ln 892, Col 86

NUMBER OF LINES IN STOCK.TXT AFTER THE ANALYSIS

Let's see some of the pastes found:

First file:

ixyIZwY5.gz contains Game related conversations+25 occurrences of a game related word

FIRST FILE RANDOMLY CHOSEN

The file:

```
*****MM - Resources*****
- BASE:
  - ETaC - Resources (Expanded Towns and Cities by missjennabee): http://www.nexusmods.com/skyrim/mods/13608/?
- FACE PARTS:
  - SV Beards version 2.0: http://shadowtigers.tumblr.com/post/135867651864/sv-beards-20
  - Lux Brows by StephieRawx: http://stephierawx.tumblr.com/post/159408387390/lux-brows-by-stephierawx-contains-20-standalone
  - Vanilla Hair Variety Plus by Omega99jp: https://www.nexusmods.com/skyrim/mods/28936/?
  - Hallgarth's Additional (Vanilla) Hair by Hallgarth: https://www.nexusmods.com/skyrim/mods/78669/?
  - Hirsute Beards by Snfkin: https://www.nexusmods.com/skyrim/mods/79442/?
  - Northborn Scars by Northborn: http://www.nexusmods.com/skyrim/mods/49279/?
  - Lovely Hairstyles by sn00p (zn00p): https://www.nexusmods.com/skyrim/mods/7403/?
  - Maevan2's eye brows by Maevan2: https://www.nexusmods.com/skyrim/mods/72825/?
  - Northborn Scars by Northborn: http://www.nexusmods.com/skyrim/mods/49279/?
  - Hirsute Beards by Snfkin: http://www.nexusmods.com/skyrim/mods/79442/?
  - Conan Hair - a Hyborian Haircut for Barbarians by Viltuska: http://www.nexusmods.com/skyrim/mods/59478/?
  - Howitzer's Hair for Men by Howitzer155: http://www.nexusmods.com/skyrim/mods/54254/?
  - Khajiit Hair by Saerileth: http://www.nexusmods.com/skyrim/mods/67873/?
  - SG Female Eyebrows by Hello Santa: http://www.nexusmods.com/skyrim/mods/35327/?
  - The Eyes Of Beauty by Gabriel Mailhot as LogRaam: http://www.nexusmods.com/skyrim/mods/13722/?
  - Lind's Human Eyes by Lind001 - LindsWorkshop: http://www.nexusmods.com/skyrim/mods/75674/?
  - Lind's Elven Eyes by Lind001 - LindsWorkshop: http://www.nexusmods.com/skyrim/mods/74948/?
  - Cherry's Eyes by CherryMods: http://www.nexusmods.com/skyrim/mods/60111/?
- FAUNA:
  - Farm Animals by Ian Joseph (Ianjoseph1986): http://www.nexusmods.com/skyrim/mods/56961/? -> tweaked
  - Varied Chickens by Jokerine: http://www.nexusmods.com/skyrim/mods/57732/? -> added to leveled lists
```

IXYZWY5.TXT

As we can see the file is a simple list of mods for skyrim. So, the module worked and found a game related file.

Another file:

`z1hRvE0c.gz` contains Game related conversations+8 occurrences of a game related word

SECOND FILE

```
wait(1)
canspirit = false
goup = 1
spiritballenergy = false
local spirit1 = false
local spirit2 = false
local sizeup = 38
local ringgo = 18
local potara = false
local potara1 = false
local potara2 = false
local potara3 = false
local potara4 = false
local hipheight = false
local idle7 = true
local idle6 = true
local idle5 = false
local noidle = false
local noidle1 = false
local fuse = false
local bigkamehameha1 = false
local bigkamehameha2 = false
local idle2 = false
local idle3 = false
local com1 = 10
local com2 = -5
local fuse1 = false
local fuse2 = false
local fuse3 = false
local fuse4 = false
local fuse5 = false
local fuse6 = false
local fuse7 = false
local fuse8 = false
local ssj4 = false
local headcolor = 0
local walk11 = true
local great = false
local size = 0
local size2 = 0
local kicharging = false
local ki = 100
local ScreenGui = Instance.new("ScreenGui")
local Gui = Instance.new("Frame")
local Frame = Instance.new("Frame")
local Frame_2 = Instance.new("Frame")
local Frame_3 = Instance.new("Frame")
```

Z1HRVE0C.TXT

We can see that this file is maybe a configuration file for a dragonball game. So, the module worked again.

CONCLUSION

In this project we had the chance to discover a very interesting framework and to see how it works.

We created a module and managed to put it in AIL. We had some issues with this part of the project because the first step to create a module was explained but not the work after that to make it work in AIL.

But we managed to do that, and our module works well but we think it can be more accurate if we change the list or the number of game words required to be detected.

We can even think other uses for this module with other type of subject.