A CIP-PSP funded pilot action
Grant agreement n°325188

Deliverable	**D3.2 Design report of experiments**
Work package Due date Submission date Revision Status of revision	WP3 Experiment Planning, Integration and Deployment M18 1.0
Responsible partner Contributors	INTECO (Angela García, Gonzalo de la Torre, Jonás Ropero, Ana Santos)
Project Number Project Acronym Project Title Start Date of Project	CIP-ICT PSP-2012-6 / 325188 ACDC Advanced Cyber Defence Centre 01/02/2013

Dissemination Level	
PU: Public	
PP: Restricted to other programme participants (including the Commission)	X
RE: Restricted to a group specified by the consortium (including the Commission)	
CO: Confidential, only for members of the consortium (including the Commission)	

Version history

Rev.	Date	Author(s)	Notes
v.0.1	21/07/2014	Angela García (INTECO), Gonzalo de la Torre (INTECO), Jonás Ropero (INTECO), Ana Santos (INTECO).	Initial version to complete and review by wp3 partners
v.1.0	31/07/2014	Antonio Pastor (TID), Beatriz Gallego-Nicasio (ATOS), Darko Perhoc (CARNET), Jochen Schoenfelder (DFN-CERT), Christian Keil (DFN-CERT), Catalin Patrascu (CERT-RO), Aleš Černivec (XLAB) Angela García (INTECO), Gonzalo de la Torre (INTECO), Jonás Ropero (INTECO), Ana Santos (INTECO).	First version (integrates all contributions from partners)

Table of contents

Table of figures

Table of tables

1. Executive summary

Along this document, it will be described the specific design for each experiment defined in ACDC: Spam Botnet, FastFlux, DDoS, Websites and Mobile.

The detailed design is defined to achieve objectives and success criteria specified in document D3.1-Planning of Experiments.

This document contains general sections that apply to all experiments and one specific section for each experiment.

General sections contain the following information:

- Section 2: Deployment and integration activities to be done by experiment participants.
- Section 8: Notification and mitigation activities to be done by ISPs, CERTs and National Support Centers in the scope of the experiment.
- Section 9: General experiment conditions that should be met for the success of the experiments.
- Section 10: Guidelines for confidence level definition and reporting.
- Section 11: Annex with examples of notification and mitigation activities.

Specific sections for experiment (sections 3 to 8) contain:

- Specific definitions and terms in the scope of the experiment.
- Flow of processes and activities to be done by different roles in the experiment. Roles are defined in document D3.1, so based on these roles, partners or participants can know which activities applies to them in each part of the experiment. Also, some activities will be associated with specific datasets (in the sharing data flow process).
- Experiment data flow diagram.
- Definition of datasets and use cases examples for the experiment.
- Metrics to provide by each participant on the experiment, depending on the role played.
- Procedure and templates to report the results.

This document does not describe specific partner technologies involved; neither lists participants in the experiment or procedures to become a participant on a specific experiment. This document details how the experiments will be executed in order to achieve the objectives defined in D3.1. Technologies and participants are identified also on D3.1.

2. Deployment and Integration activities for all experiments

This section defines preliminary process and activities to execute by all experiment participants: deployment and integration in the ACDC system architecture.
As defined in D2.3 document, there are different deployment approaches on ACDC: as proprietary services sharing data, as open solutions that are deployed on third party networks, as final end-user deployments, etc. For any of the deployment model used, as indicated in D3.1 - section 2.8, all partners and tools that are going to contribute to any ACDC experiment sharing data (sending and/or retrieving), must be integrated with the Centralized Data Clearing House and publish its participation an contribution through the Community Portal. Also, for the datasets that each participant will send, data sharing policies should be applied using data sharing procedures defined also in the Community Portal.

In any case, data sent and retrieved from ACDC on each experiment must be managed by each participant complying with data protection and privacy laws following recommendations and requirements in D1.8.1 Legal document of ACDC and the specifications given for the different types of solutions.

One of the main activities of this preliminary phase is that each participant must define the specific datasets to share and the schemata to use for sharing data through ACDC. To do this, general data schemata defined in D1.7.2 document: Data formats Specification, must be used. Along this document, on the specific sections for each experiment, main experiment datasets are described and examples of sharing this information is given using the Global ACDC data Schemata.

The following table defines the deployment and integration process and activities to be done by participants:

Process		Description	Activities		Role[1]	Input Info	Output Info
1	**Tool owner's deployment**	Deploy the tool in the tool owner infrastructure. It should be tested, configured and ready to be run on the experiment.	1.1	Install the tool in the tool owner network.	Tool Owner	Tool SW package. Installation & configuration manual. Input data sources User/operation manual.	Send notification to the parts involved in the experiment, that the tool is up and running and ready to integrate.
			1.2	Check for correct operation.			
			1.3	Run the tool.			

[1] Roles are defined in Document D3.1

2	**Partner's Deployment**	Deploy the tool in other partners infrastructures with the aim of obtain the maximum number of input data sources in order to increase detections.	2.1	Find partner networks in which deploy the tool.	Tool Operator & Tool Owner	Tool SW package. Installation & configuration manual. Input data sources User/operation manual.	Send notification to the parts involved in the experiment, that the tool is up and running and ready to integrate.
			2.2	Provide the tool and the documentation to the partner.			
			2.3	Install the tool in the partner network.			
			2.4	Check for correct operation.			
			2.5	Run the tool.			
3	**End User's Deployment**	Distribute and disseminate end-users tools to be installed on end-users devices.	3.1	Deliver the tool to the National Support Centers.	Tool Owner & NSCs & ACDC Dissemination Team	Tool documentation. Tool package.	Reports sent by National Support Centers and Tool Owners. Dissemination Activities results.
			3.2	Publish and disseminate the tool through National Support Centers and/or other channels to find end users who will download and install the tool.			
4	**CCH Integration**	Integrate the tool with the CCH.	4.1	For each Tool, define the datasets and the schemas that will be used to share the information with the CCH.	Tool Owner & Tool Operator Network owners CERTs NSCs	CCH API Documentation. Community Platform (CP) procedures. General ACDC Data Schemata (Document D1.7.2 : ACDC Data Schemata specification). Specific Datasets for the experiments: SPAM WEBSITES FASTFLUX DDOS MOBILE	Specific Datasets Schemas to be used for each participant published in the CP (if differ or extends the Global ACDC schemata)..
			4.2	Put in place the necessary local processes for each tool in order to be ready to start sending and/or collecting information with the CCH. The integration test also must be done by partners involved in the processes of mitigation and notification: ISPs, NSCs and CERTs who have to retrieve information from CCH.			
			4.3	Test the integration.			
			4.4	Publish the Dataset Schema to be used by the Tool in the Community Platform (CP) in order to be known by all partners (if differ or extends the Global ACDC schemata).			

Table 1 - Deployment and integration process and activities

3. SPAM experiment design

The design of this experiment is defined to achieve objectives detailed in section 3.1 of document D3.1-Planning of Experiments.

3.1. Specific definitions for SPAM experiment

Taking in mind that this experiment is focused on detection an analysis of spam messages to detect and mitigate spam-botnet elements, the following terms are defined in the scope of the experiment:

SPAMBOT	
DEFINITION	The automated program or piece of malware that sends spam from compromised devices. In this experiment a spambot is identified by at least a public IP address and the TIMESTAMP of the detection of the spambot activity.
DETECTION	Different technologies and criteria can be used to identify spambots in the experiment: - Spambot detection is done by suspect behaviour observation. Suspicious information can be detected into the header and protocol. Some honeypots or feedback loops are used for spambot observation and detection. - Spam messages where the server IP belongs to dynamic IP address spaces and not from mail servers (that normally has fixed IP addresses). - Set of dynamic IPs participating in the same spam campaign. - IPs of devices connecting to a C&C server belonging to a spam botnet. - Etc.
DATASET	Spam attack dataset

Table 2 - SPAM definition - SPAMBOT

SPAM CAMPAIGN	
DEFINITION	For this experiment, a spam campaign is defined as a message sent by spambots to multiple users in a period of time to achieve illegal activity. The campaign message can contain additional suspicious elements that must be identified and analyzed inside the experiment: like attached files and URLs.
DETECTION	Different technologies and criteria can be used to identify spam campaigns in the experiment: - Spam messages with same subject or specific patterns defined. - The IPs addresses involved in sending those messages are considered to be spambots, if matching the spambot detection criteria. - Different campaign data can be correlated in order to find spambots that belong to the same botnet or to identify another detection rules. The spam campaign must be described on the datasets in order to be able to advice end-users potentially affected through the National Support Center Websites of the project.
DATASET	Spam campaign dataset

Table 3 - SPAM definition - SPAM CAMPAIGN

C&C SERVER	
DEFINITION	A C&C server member of a botnet focused on sending spam. In the scope of the experiment, also a C&C server found as a result of the analysis of the different elements detected on the spam experiment.
DETECTION	C&C servers can be detected from the analysis of: - Spambot reversing analysis - Malware found in spam messages - Analysis of URLs found in spam messages - Campaign correlation and analysis
DATASET	Spam C&C dataset

Table 4 - SPAM Definition - C&C SERVER

SPAM SUSPICIOUS ELEMENTS	
DEFINITION	URLs or any type of file found in spam messages and consider suspicious. It must be further analysed.
DETECTION	Processing of spam messages. Both types of elements within this category should be analyzed by a URL analyzer or a malware analyzer in the scope of the experiment.
DATASET	Spam suspicious elements dataset

Table 5 - SPAM Definition - SPAM SUSPICIOUS ELEMENTS

SPAM MALICIOUS ELEMENTS	
DEFINITION	URLs or any type of file found in spam messages and consider as malicious.
DETECTION	Processing of spam messages and URL or malware analysis must be done to report the element as malicious.
DATASET	Spam malicious elements dataset

Table 6 - SPAM Definition - SPAM MALICIOUS ELEMENTS

BOT	
DEFINITION	The automated program or piece of malware that sends spam from compromised devices. In this experiment is a subgroup of the spambot and it is also identified by at least a public IP address and the TIMESTAMP of the detection of the spambot activity. To identify a Bot it is not necessary to observe it actively participating in an attack.
DETECTION	Different technologies and criteria can be used to identify bots in the experiment: - Extracted from a sinkholing or similar techniques. - IPs of devices connecting to a C&C server belonging to a spam botnet. - Etc.
DATASET	Bot dataset

Table 7 - SPAM Definition - BOT

3.1.1. Confidence Level of the information

Independently of the type of element or incident identified, each report shared through the Central Clearing House (CCH) must indicate the level of veracity of the information (through the ***confidence_level*** parameter on the datasets). This is very important for the notification and mitigation part of the experiment.

Common criteria can be applied following guidelines in section 10 of this document.

3.2. Experiment processes and activities

3.2.1. Detection and analysis

The following table details the process and activities to execute along the experiment time (process 1 to 4 are the same for all experiment as defined in section 2 of this document). **This table covers detection and analysis activities**.
Not all the activities must be performed by the role identified. Inside the experiment each participant defines the scope of its role and therefore the scope of the actions to execute.

Process		Description	Activities		Role[2]	Input Info	Output Info
SP1	**Tool detection phase: Collecting data from CCH**	OPTIONAL: Collect information from the CCH useful to feed systems spam-botnet sensors in order to increase number and quality detections. This process is a constant task through the experiment.	SP1.1	Request the necessary information needed based on the datasets available on the CCH.	Tool Operator	Datasets available in CCH.	New detection rules for sensors
			SP1.2	Feed the detection tool with the information obtained.			
SP2	**Tool detection phase: Spam Email harvesting**	Through the spamtraps tools, end user tools and network traffic sensors, harvesting spam data. This data will be used to detect and identify spam bots, infection channels and malware, and to obtain valuable data for statistics.	SP2.1	Collect spam messages through spamtraps and/or end-users tools.	Tool Operator	Honey Tokens	SPAM messages (body, header of the email, subject, attachments, email server logs, malware hash, URLs embedded in spam body…)
			SP2.2	Capture SMTP traffic.	Network Owner and Tool Operator	Network traffic	IPs Information about spammers
SP3	**Tool detection & analysis phase: Classification,**	The spam messages/smpt traffic collected in the previous process must be analyzed and classified by sensors to identify spam-botnet	SP3.1	Identify spambots	Tool Operator	Information collected from the process SP1 and SP2.	Reports with the data obtained (based on dataset schemata defined).
			SP3.2	Identify spam-botnet campaigns	Tool Operator		

[2] Roles are defined in Document D3.1

	analysis and identification of spam-botnets related elements.	elements and/or infection channels.	SP3.3	Identify malware in spam	Tool Operator	SPAM datasets Schemas definition: • Spam attack dataset. • Spam campaign dataset. • Spam C&C dataset. • Spam suspicious elements dataset. • Spam malicious elements dataset. • Spam botnet dataset • Bot dataset	
			SP3.4	Identify malicious URL in spam	Tool Operator		
			SP3.5	Identify C&C servers	Tool Operator		
SP4	**Data Correlation**	Correlation of data in order to increase spam-botnet detections and new rules and events	SP4.1	Correlate the data detected and shared by all partners.	Tool Operator	Data extracted from process SP3.	Reports with the data obtained (based on dataset schemata defined).
SP5	**Delivery data to CCH**	Delivery to the CCH all data and information collected in previous phases.	SP5.1	Send information obtained to the CCH. See response time guidelines for the experiment.	Tool Operator	Information collected and correlated (if apply) from the process SP3 and SP4: Reports with the data obtained (based on dataset schemata defined).	
SP6	**Periodic Control Report (Detection & Analysis report by tool)**	Generate a periodic report in order to keep track of the experiment with the information obtained during the experiment detection phase. It must be sent to the experiment coordinator with the frequency stipulated.	SP6.1	Generate the report following the template supplied by leaders.	Tool Operator	Data from process SP1, SP2, SP3, SP4 & SP5 Periodic Report Template (Detection & Analysis phase)	Periodic Control Report (Detection & Analysis report by tool)
			SP6.2	Send the report to the experiment leaders (INTECO & CARNET).			
SP7	**CCH Monthly Report**	Periodically, generate a report with global SPAM metrics.	SP7.1	Generate a report with metrics about the information received and collected in the CCH during the last month regarding SPAM.	CCH Operator	Information in the CCH. Inputs and outputs requests by partners.	CCH Report

			SP7.2	Send the report to the experiment leaders (INTECO & CARNET).			

Table 8 - SPAM - process Detection and analysis

3.2.2. Notification and mitigation

Notification and mitigation activities are very similar on design along the different experiments, so these activities are explained for all in section 8 of this document.

Specific to SPAM experiment is the identification of spam campaigns, and so on, the advertising of this campaigns through the National Support Centers (NSCs). This means that NSCs, in the scope of the experiment must:

- Retrieve spam campaigns from CCH
- Analyze which ones are affecting to users of its country (for example a phishing campaign to a National Bank)
- In case positive, generate the content and advertise about it through the NSC web portal (See success criteria defined in D3.1).

An example of this can be found in Annex I.

3.2.3. Response times

Some activities of the experiment require maximum response times in order the whole process to be effective. This response times are defined for all experiments in section 9 of this document.

3.3. Experiment Data Flow Diagram

The following diagram shows the dataset flow between different components along the different phases or process of the experiment:

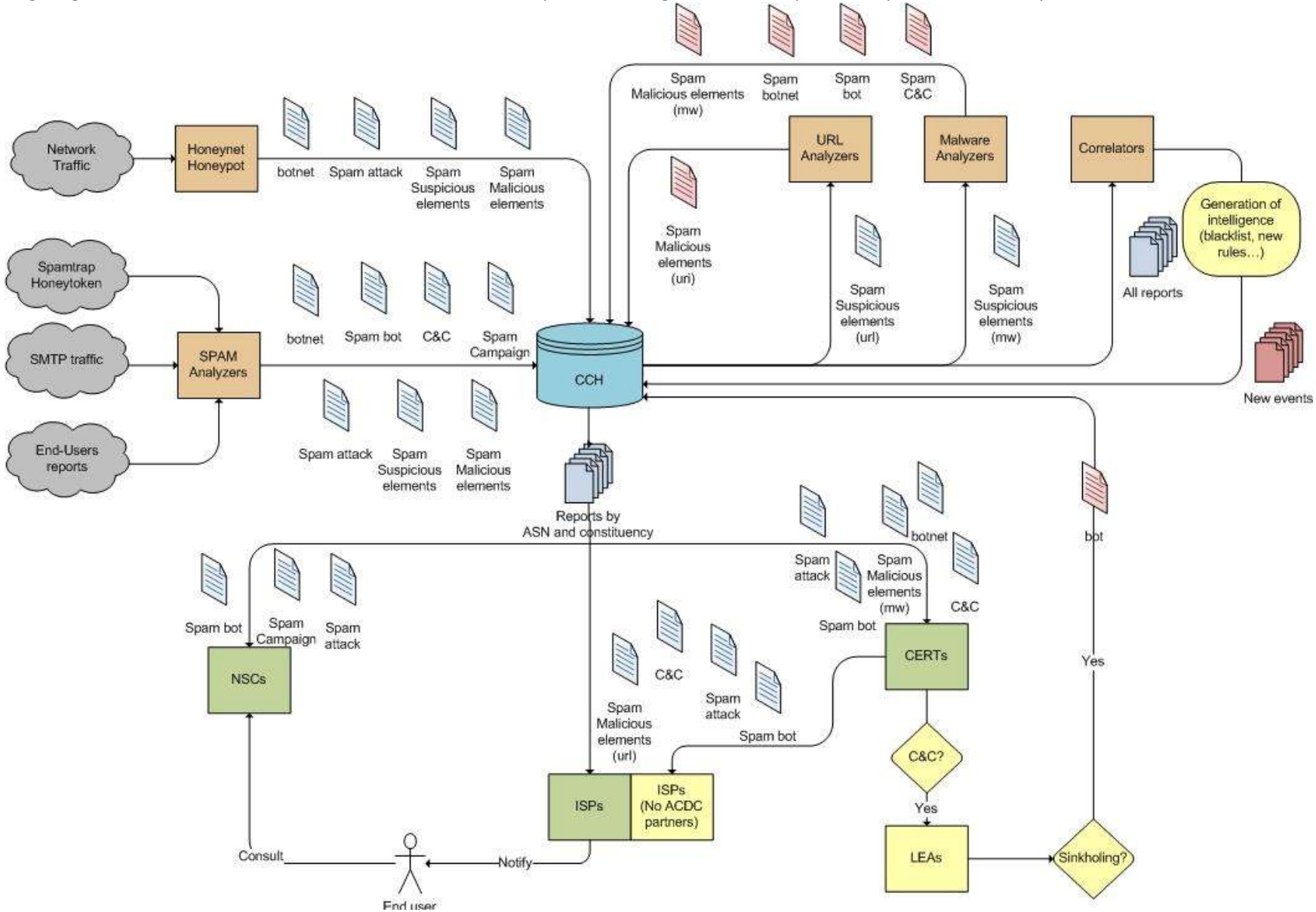

Figure 1 - SPAM experiment data flow

3.4. Datasets definition for SPAM experiment

Based on the specific spam elements to detect and analyse in the scope of the experiment (section 3.1), and on the data schemata defined at the Document D1.7.2 Data Formats Specification, the following datasets has been defined:

- Spam attack (common spam bot).
- Spam campaign.
- Spam C&C.
- Spam suspicious elements.
- Spam malicious elements.
- Spam botnet.
- Bot.

The fields defined in each dataset are the minimum data for the experiments but they could be extended and any other field can be added by participants.

Extended datasets used must be defined and published through the Community Portal in order to be known by all participants on the experiment.

The following tables contains, for each field defined: a functional description, the field name, the type, and its obligation. In fields with multiple possible values there are specified only those that are involved in this experiment. It also includes some optional fields that are not necessary to send if they are not known.

3.4.1. Spam attack dataset

The following dataset represent the minimum specific data that must be sent for each spam bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.attack	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object.	source_key	string enum: ip	False
IP of the bot.	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with	confidence_level	number minimum: 0.0 maximum: 1.0	False

0.0 being unreliable and 1.0 being verified to be accurate.			
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the attack.	report_subcategory	enum: abuse	False
The RFC 790 decimal internet protocol number of the attack connection.	ip_protocol_number	integer minimum: 0 maximum: 255	False
The IP version of the attack connection.	ip_version	integer enum: 4, 6	False
The botnet the attack can be attributed to (if apply).	botnet	string	True
The IP of the spam bot.	src_ip_v4	string format: ipv4	False
	src_ip_v6	string format: ipv6	False
The destination port of the attack connection.	dst_port	integer	False
The subject of the associated campaign to this attack.	subject_text	string	False

Table 9 - SPAM dataset attack

3.4.2. *Spam campaign dataset*

The following dataset represent the minimum specific data that must be sent for each spam campaign.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.spam_campaign	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an email subject.	source_key	string enum: subject	False
The subject, body or header of the campaign.	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being	confidence_level	number minimum: 0.0 maximum: 1.0	False

verified to be accurate.			
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet associated to the campaign.	botnet	string	True
Additional data for the observation, as the email body and header anonymized and also a brief description of the criteria used to define the campaign.	additional_data	object	True
The filename of the malicious attachment used in this campaign.	sample_filename	string	True
The malicious uri (can be more than one) associated with this campaign.	malicious_uri	string format: uri	True
A general description of the campaign in order to can identify it; social engineering used, type of attachments…	description	string	True

Table 10 - SPAM dataset Campaign

3.4.3. Spam C&C dataset

The following dataset represent the minimum specific data that must be sent for each spam C&C.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.c2_server	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object.	source_key	string enum: ip	False
The IP address of the C&C server.	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0	confidence_level	number minimum: 0.0 maximum: 1.0	False

and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.			
The version number of the data format used for the report.	version	integer enum: 1	False
The control channel used by the C2.	report_subcategory	string enum: http, irc, other	False
The botnet associated to the C&C.	botnet	string	True

Table 11 - SPAM dataset C&C

3.4.4. Spam suspicious elements dataset

The following dataset represent the minimum specific data that must be sent for each spam suspicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

3.4.4.1. Suspicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious content at the uri.	report_subcategory	string enum: exploit,	False

		malware, phishing, other	
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 12 - SPAM dataset suspicious uri

3.4.4.2. Suspicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True
The binary of the sample encoded in base 64.	sample_b64	string	True

Table 13 - SPAM dataset suspicious malware

3.4.5. Spam malicious elements dataset

The following dataset represent the minimum specific data that must be sent for each spam malicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

3.4.5.1. Malicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious content at the uri.	report_subcategory	string enum: exploit, malware, phishing, other.	False
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 14 - SPAM dataset malicious uri

3.4.5.2. Malicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True
The binary of the sample encoded in base 64.	sample_b64	string	True

Table 15 - SPAM dataset malicious malware

3.4.6. Spam botnet dataset

The following dataset represent the minimum specific data that must be sent for each spam botnet.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.botnet	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one	report_type	string	True

sentence.			
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: botnet	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The version number of the data format used for the report.	version	integer enum: 1	True
The category of the botnet	report_subcategory	string enum: c2, p2p, other	True

Table 16 - SPAM dataset botnet

3.4.7. Bot dataset

The following dataset represent the minimum specific data that must be sent for each bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.bot	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The timestamp when the reported observation took place.	timestamp	string format: date-time	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: ip	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	True
The version number of the data format used for the report.	version	integer enum: 1	True
The botnet the bot is attributed to.	botnet	string	True
The IP of the C&C where the bot is involved.	c2_ip_v4	string format: ipv4	False
	c2_ip_v6	string format: ipv6	False

Table 17 - SPAM dataset bot

3.5. *Dataset examples*

A functional example of the main dataset flows for this experiment is:

Scenario 1:

A campaign has been detected by some partner/sensor.
This campaign is associated with the following name: campaign_medusa

The campaign contains the following elements:

- *Subject*: Medusa Corp Looking For People
- A pdf *attachment* (medusacorp.pdf)

Sensors have considered that it's a campaign by detecting several spambots sending the same mail in a short period of time:

- spambot1medusa, 10.10.10.1
- spambot2medusa, 10.10.20.1
- spambot3medusa, 10.10.30.1

After a previous analysis, it's confirmed that the attachment is a malicious one.

Dataset sent for scenario 1:

The datasets that take place on this scenario are: **Campaign, attack and malware**:

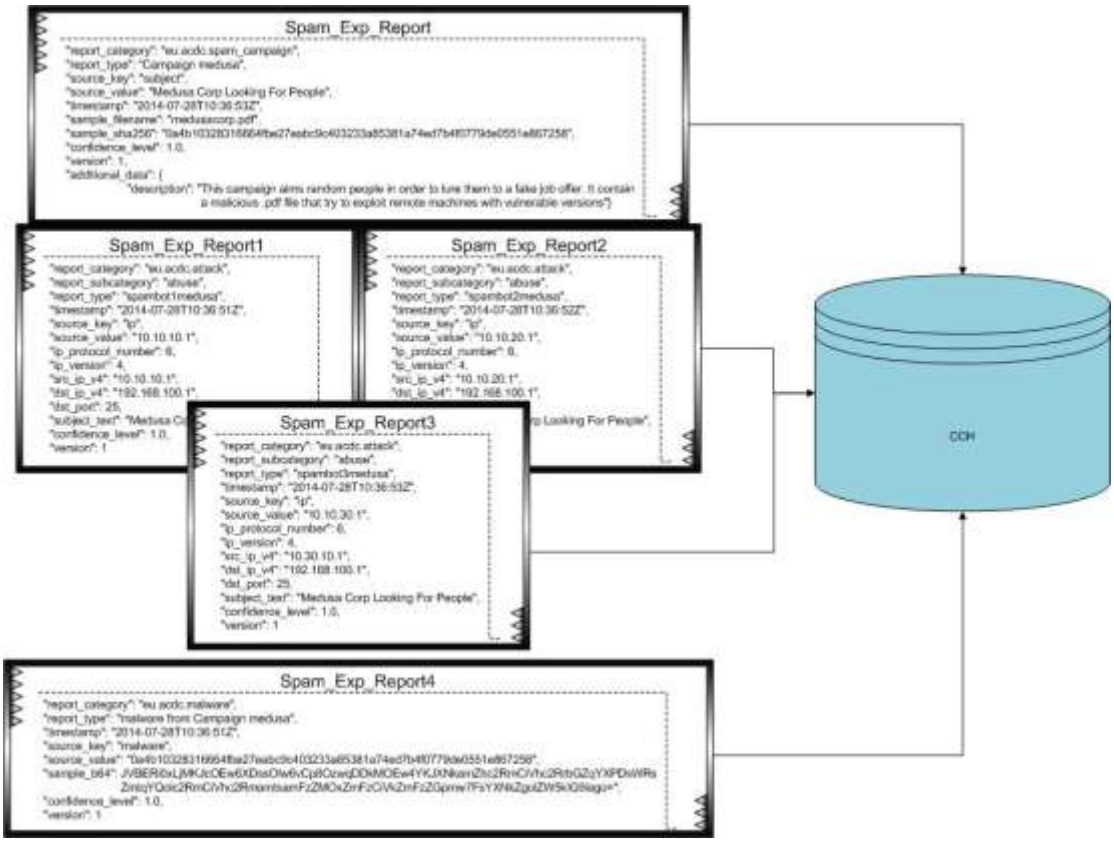

Figure 2 - SPAM dataset example

3.6. Metrics

Depending of the role of each participant the following metrics must be reported in the intermediate reports through the Templates defined by experiment leaders:

Experiment Phase	Metric	Description	Classified by (if applies)
General	Partners	Number of partners participating on the experiment	• Type of organization • Role in the experiment • Technologies
Deployment & Integration	Tools	Number of tools contributing to the experiment	• Number of deployments • Contribution type
Detection & analysis	Spam volume	Number of spam messages detected and analyzed	• Detection Tool • ASN • Country
	IPs sending spam	Number of total IPs addresses identified sending SPAM	• Detection Tool • ASN • Country
	Spambots	Number of spambots identified (IP+TS)	• Detection Tool • ASN • Country • Per campaign identified
	C&C	Number of C&C servers identified on the experiment scope	• Detection Tool • ASN • Country
	Campaigns	Number of campaigns identified	• Total • Number that distribute MW in attachment • Number that distribute malicious URL (type if possible) • Size (number of messages identified) • Localization (countries affected by spambots involved)
	URLs in SPAM	Number of URLs found in SPAM	• Total • Total analyzed • Total Malicious • % sent by spambots • Total Malicious per ASN • Total malicious per TLD • Type of malicious activity
	MW in spam	Number of attachments sample found in SPAM	• Total • Total analyzed • Total Malicious • % sent by spambots

			• Type of malicious activity • Per campaign identified
	Botnets	Number of different botnets detected	• Total
Data Storage	Total SPAM Reports in CCH	Number of reports sent to CCH related to spam	• Total accumulated • Per day/week • Per Tool/Partner
Distribution for notification & mitigation purposes	SPAM Reports retrieved	Number of SPAM reports retrieved for analysis, notification and mitigation	• Total per partner ASNs (depending the ISP, network owner or CERT constituency) • Per type of element retrieved • For ISPs: Classification per type of network affected (mobile or fixed)
Notification	Notifications	Notifications sent to end users and processes activated with LEAs.	• Total • ASN • Type of element • Sent to: end-user, ISP, LEA.
Mitigation	Campaigns published	Number of Spam Campaigns advisories on NSCs.	• NSC • Month
	SPAM prevention and mitigation contents/tools	Number of visits/downloads to SPAM contents/tools in NSCs	• NSC • Month

Table 18 - SPAM metrics

3.7. Reports

During the execution period of the experiment, each participant must complete and send a periodic report (PR) to experiment leaders.

Depending the role in the experiment and the tools operated, this report must contain:

- The metrics
- Incidents or problems during the period
- Specific considerations and conclusions

By default, the PR will be sent weekly, unless a different periodicity could be needed.

Experiment leaders will send a Periodic Report Template per experiment to each participant.

The report Template: **ACDC_EXP_SPAM_PR_template.xls** (annexed to this document) will be available also through the Community Portal website.

A final and global report will be developed by experiment leaders. Main conclusions and results will be published on the CP website at the end of each experiment.

4. WEBSITES experiment design

The design of this experiment is defined to achieve objectives detailed in section 4.1 of document D3.1-Planning of Experiments.

4.1. Specific definitions for WEBSITE experiment

Taking in mind that this experiment is focused on detection of malicious websites used by botnets or botnets used to attack and infect websites, the following terms are defined in the scope of the experiment:

WEBSITE	
DEFINITION	The identifier (URI) of the website that is or can be involved in botnet activities.
CLASSIFICATION	In the scope of the experiment websites can be classified by: **Suspicious**[3] => need to be analyzed **Malicious** => develop some type of malicious activity **Vulnerable** => websites vulnerable that can be compromised Malicious websites can be classified also by: - Malware - Exploit - Phishing - Other
DETECTION	Different technologies and criteria can be used to identify vulnerable and malicious websites: - Proactive scanning of websites - Reports by third parties (end-users or collaborators) - Honey nets. - Attacks identified originated from websites - ...
DATASETS	Website suspicious elements dataset Website malicious elements dataset Website vulnerable dataset

Table 19 - WEBSITE definition - WEBSITE

MALWARE (DISTRIBUTED ACROSS WEBSITES)	
DEFINITION	Files or code found in a WEBSITE suspicious to be malware
ANALYSIS	If the file is reported as suspicious, it must be analyzed by a Malware analyzer in the scope of the experiment. If the file is reported as malicious it must be described as indicated on the specific dataset.
DATASET	Website malicious elements dataset

Table 20 - WEBSITE definition - MALWARE

[3] Is consider as suspicious if there are enough evidences of it is doing any malicious activity but it needs a deeper analysis to confirm it.

WEBSITE BOT	
DEFINITION	Compromised or malicious website controlled by a botnet to perform specific illegal activities, for example: malware distribution, phishing, etc. Also, the automated program or piece of malware installed on end-user devices that search and scans legitimate websites with the objective of compromise them, for example. By this way the website can contribute to the botnet activity. In the experiment a website bot is identified by at least a public IP address and the TIMESTAMP of the detection of the malicious activity. To identify a Bot it is not necessary to observe it actively participating in an attack.
DETECTION	Different technologies and criteria can be used to identify website bots in the experiment: - Sinkholing. - Malware analysis. - IDSs. - Website analyzers. - Etc.
DATASET	Website bot dataset

Table 21 - WEBSITE definition - BOT

WEBSITE ATTACK	
DEFINITION	Actions carry out against a website in order to obtain unauthorized access to it or realize any other malicious action, like try to upload a malware. In the experiment a website attack is identified by at least a public IP address of the system performing the attack and the TIMESTAMP of the detection of the malicious activity.
DETECTION	Different technologies and criteria can be used to identify website bots in the experiment: - Honeynets. - IDSs. - Website analyzers. - Etc.
DATASET	Website attack dataset

Table 22 - WEBSITE definition - ATTACK

C&C SERVER	
DEFINITION	A C&C server member of a botnet used to develop malicious websites activities (like malware distribution, phishing, fraud or any illegal activity). In the scope of the experiment, also a C&C server found as a result of the analysis of the different elements detected on the website experiment.
DETECTION	C&C servers can be detected from the analysis of: - Websites - Malware found in websites - Correlation activities.
DATASET	Website C&C dataset

Table 23 - WEBSITE definition - C&C SERVER

4.1.1. Confidence Level of the information

Independently of the type of element or incident identified, each report shared through the Central Clearing House (CCH) must indicate the level of veracity of the information (through the ***confidence_level*** parameter on the datasets). This is very important for the notification and mitigation part of the experiment.

Common criteria can be applied following guidelines in section 10 of this document.

4.2. Experiment processes and activities

4.2.1. Detection and analysis

The following table details the process and activities to execute along the experiment time (process 1 to 4 are the same for all experiment as defined in section 2). **This table covers detection and analysis activities**.
Not all the activities must be performed by the role identified. Inside the experiment each participant defines the scope of its role and therefore the scope of the actions to execute.

Specific Processes		Description	Activities		Role[4]	Input Info	Output Info
WB1	**Tool detection phase: Collecting data from CCH**	OPTIONAL: Collect information from the CCH needed to feed website sensors in order to increase number and quality of detections. This process is a recurring task throughout the experiment.	**WB1.1**	Request the necessary information needed (if apply).	Tool Operator	Datasets available in CCH.	New detection rules
			WB1.2	Feed the detection tool with the obtained information.			
WB2	**Tool detection phase: Detect malicious websites**	Check, using the website sensor tools and end user tools, whether a site is using any malicious or suspicious techniques and classify it. This data will be used to detect and identify website bots, infection channels and malware, and to obtain valuable data for statistics.	**WB2.1**	Check if the website gets malicious or suspicious content.	Tool Operator	WebSites data. WEBSITES datasets Schemas definition: • Website attack. • Website suspicious elements. • Website malicious elements. • Website vulnerable. • Website C&C. • Website botnet. • Website bot.	Reports with the data obtained (based on dataset schemata defined)
			WB2.2	Classify the malicious website: - Malware - Malicious content - Fraud - Access/intrusion - Other	Tool Operator		
WB3	**Tool detection**	Through the website tools,	**WB3.1**	Detect and collect attack data to	Tool	Information collected	Reports with the

[4] Roles are defined in Document D3.1

	phase: Detect and analysis of attacks against a website.	honeypots/honeynets and end user tools detect any active botnet focused on attacks to websites and also detect the bots related to the botnet and the evidences of the attack.		websites.		Operator	by the tool. WEBSITES datasets Schemas definition.	data obtained (based on dataset schemata defined)
WB4	**Tool detection phase: Analysis of the malicious websites and their related elements**	The malicious website discovered during the previous process must be analyzed to identify website-botnet elements and/or infection channels.	**WB4.1**	Analyze the website extracted data.	Analysis of vulnerabilities Analysis of redirections (This action implies to restart the process) Analysis of binaries downloaded (malware) Analysis of URLs within the website (This action implies to restart the process)	Tool Operator	Information collected from the processes WB2 and WB3. WEBSITES datasets Schemas definition.	Reports with the data obtained (based on dataset schemata defined).
WB5	**Data correlation**	Correlation of data in order to increase malicious website detections, botnet detections, new malware indicators.	**WB5.1**	Correlate the data detected and shared by all partners.		Tool Operator	Data extracted from process WB2, WB3 & WB4.	Reports with the data obtained (based on dataset schemata defined)
WB6	**Delivery data to CCH**	Delivery to the CCH all data and information collected in previous phases.	**WB6.1**	Send information obtained to the CCH.		Tool Operator	Information collected and correlated (if apply) from the process WB2, WB3, WB4 & WB5 Reports with the data obtained (based on dataset schemata defined).	
WB7	**Periodic Control Report (Detection & Analysis report**	Generate a periodic report in order to keep track of the experiment with the information obtained during the	**WB7.1**	Generate the report following the template supplied by leaders.		Tool Operator	Data from process WB1, WB2, WB3, WB4, WB5 & WB6.	Periodic Control Report (Detection & Analysis report

	by tool)	experiment detection phase. It must be sent to the experiment coordinator at stipulated intervals.	**WB7.2**	Send the report to the experiment leaders (INTECO & CERT-RO).		Periodic Report Template (Detection & Analysis phase).	by tool).
WB8	**CCH Monthly Report**	Periodically, generate a report with global Websites metrics.	**WB8.1**	Generate a report with metrics containing the information received, analyzed and collected from the CCH during the last month regarding Websites.	CCH Operator	Information in the CCH. Inputs and outputs requests by partners.	CCH Report
			WB8.2	Send report to the experiment leaders.			

Table 24 - WEBSITE - Process detection and analysis

4.2.2. Notification and mitigation

Notification and mitigation activities are very similar on design along the different experiments, so these activities are explained for all in section 8 of this document.

Specific to WEBSITES experiment is the analysis of websites incidents and/or types of attacks that can be performed through websites. With this information, new contents and advisories can be developed and published through NSCs to help webmasters to protect and prevent website attacks.
Also CERTs can notify website hosting companies and webmasters about malicious activities found on their websites. An example of this can be found in Annex I.

4.2.3. Response times

Some activities of the experiment require maximum response times in order the whole process to be effective. This response times are defined for all experiments in section 9 of this document.

4.3. Experiment Data Flow Diagram

The following diagram shows the dataset flow between roles along the different phases or process of the experiment:

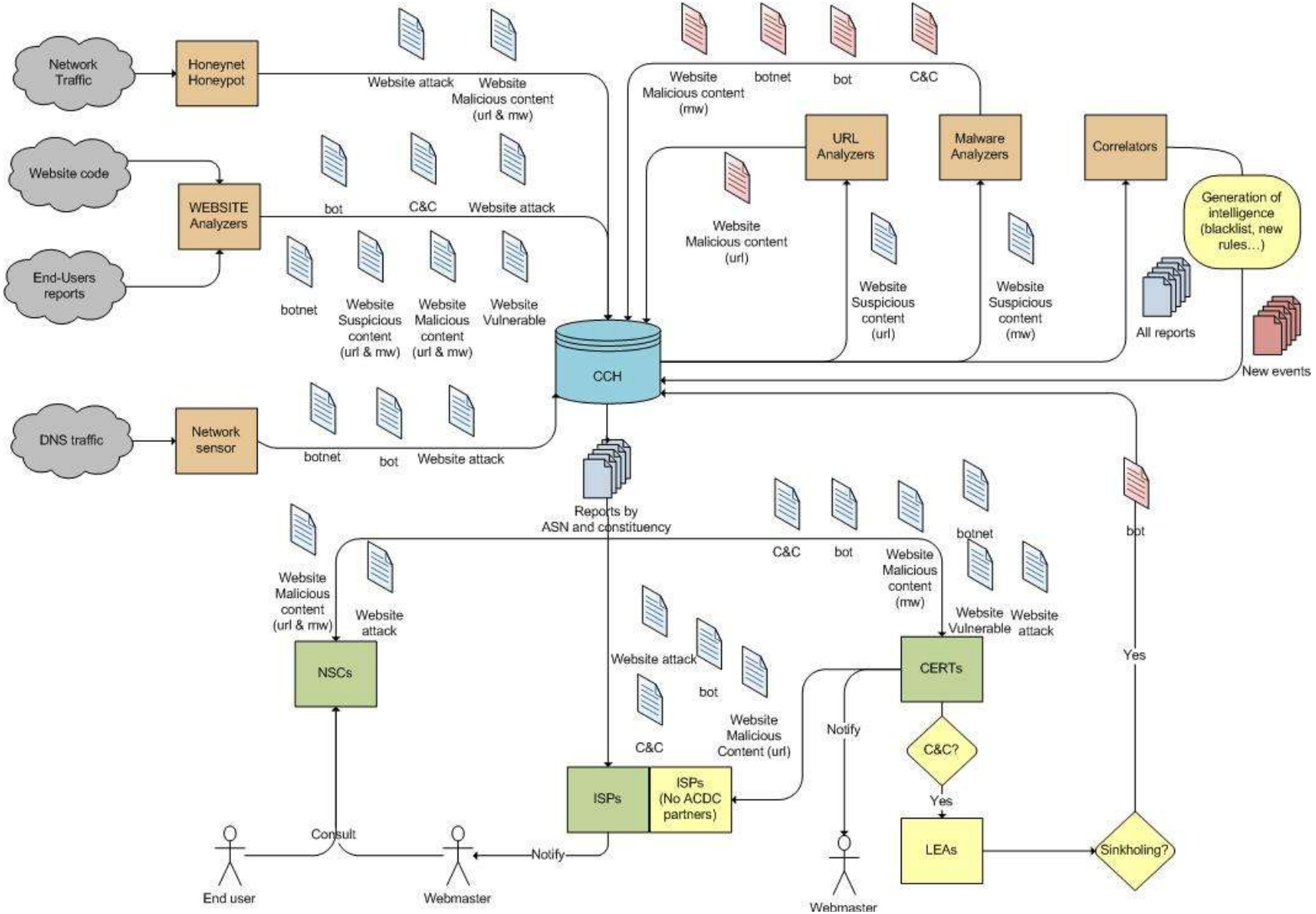

Figure 3 - WEBSITE experiment data flow

4.4. Datasets definition for WEBSITE experiment

Based on the specific spam elements to detect and analyse in the scope of the experiment (section 4.1), and on the data schemata defined at the Document D1.7.2 Data Formats Specification, the following datasets has been defined:

- Website attack.
- Website suspicious elements.
- Website malicious elements.
- Website vulnerable.
- Website C&C.
- Website bot.
- Website botnet.

The fields defined in each dataset are the minimum data for the experiments but they could be extended and any other field can be added by participants.

Extended datasets used must be defined and published through the Community Portal in order to be known by all participants on the experiment.

The following tables contains, for each field defined: a functional description, the field name, the type, and its obligation. In fields with multiple possible values there are specified only those that are involved in this experiment. It also includes some optional fields that are not necessary to send if they are not known.

4.4.1. Website attack dataset

The following dataset represent the minimum specific data that must be sent for each website attack.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.attack	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object	source_key	string enum: ip	False
IP of the system performing the attack.	source_value	string	False
The level of confidence put into	confidence_level	number	False

the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.		minimum: 0.0 maximum: 1.0	
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the attack performed.	report_subcategory	String enum: abuse, compromise, data, login, malware, other	False
The RFC 790 decimal internet protocol number of the attack connection.	ip_protocol_number	integer minimum: 0 maximum: 255	False
The IP version of the attack connection.	ip_version	integer enum: 4,6	False
The botnet the attack can be attributed to	botnet	string	True
The source IP of the attack connection. This is always the IP of the attacking system. This field equals source_value.	src_ip_v4	string format: ipv4	True
	src_ip_v6	string format: ipv6	True
The destination port of the attack connection.	dst_port	integer	False

Table 25 - WEBSITE dataset attack

4.4.2. Website C&C dataset

The following dataset represent the minimum specific data that must be sent for each website C&C.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.c2_server	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object.	source_key	string enum: ip	False
The IP address of the C&C server.	source_value	string	False
The level of confidence put	confidence_level	number	False

into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.		minimum: 0.0 maximum: 1.0	
The version number of the data format used for the report.	version	integer enum: 1	False
The control channel used by the C2.	report_subcategory	string enum: http, irc, other	False
The botnet associated to the C&C.	botnet	string	True

Table 26 - WEBSITE dataset C&C

4.4.3. Website suspicious elements dataset.

The following dataset represent the minimum specific data that must be sent for each website suspicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

4.4.3.1. Suspicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious	report_subcategory	string	False

content at the uri.		enum: exploit, malware, phishing, other	
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 27 - WEBSITE dataset suspicious uri

4.4.3.2. Suspicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True
The binary of the sample encoded in base 64.	sample_b64	string	True

Table 28 - WEBSITE dataset suspicious malware

4.4.4. Website malicious elements dataset.

The following dataset represent the minimum specific data that must be sent for each website malicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

4.4.4.1. Malicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious content at the uri.	report_subcategory	string enum: exploit, malware, phishing, other.	False
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 29 - WEBSITE dataset malicious uri

4.4.4.2. Malicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True
The binary of the sample encoded in base 64.	sample_b64	string	True

Table 30 - WEBSITE dataset malicious malware

4.4.5. Website vulnerable dataset

The following dataset represent the minimum specific data that must be sent for each website vulnerable.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.vulnerable_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of	report_type	string	False

thumb this should not be longer than one sentence.			
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: a URI.	source_key	string enum: uri	False
The uri to the vulnerable resource.	source_value	string format: uri	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.1 maximum: 1.0	False
The version number of the data format used for the report.	version	integer enum: 1	False
An array of objects describing vulnerabilities discovered at the vulnerable URI.	vulnerabilities	array items: object(identifier scheme, vulnerability identifier)	False

Table 31 - WEBSITE dataset vulnerable

4.4.6. Website botnet dataset

The following dataset represent the minimum specific data that must be sent for each website botnet

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.botnet	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: botnet	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The version number of the data format used for the report.	version	integer enum: 1	True
The category of the botnet	report_subcategory	string enum: c2, p2p, other	True

Table 32 - WEBSITE dataset botnet

4.4.7. Website bot dataset

The following dataset represent the minimum specific data that must be sent for each website bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.bot	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The timestamp when the reported observation took place.	timestamp	string format: date-time	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: IP	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	True
The version number of the data format used for the report.	version	integer enum: 1	True
The botnet the bot is attributed to.	botnet	string	True
The IP of the C&C where the bot is involved.	c2_ip_v4	string format: ipv4	False
	c2_ip_v6	string format: ipv6	False

Table 33 - WEBSITE dataset bot

4.5. Dataset examples

Two functional examples of the main dataset flows for this experiment are:

Scenario 1:

A vulnerable website has been detected by some partner/sensor.
The website content a Cross-Site Request Forgery (CSRF) identified as CWE-352.
This vulnerability has been identified before in this CMS and it has a CVE associaciated, CVE-2012-1936.

Dataset sent for scenario 1:

The dataset that take place on this scenario is: **Vulnerable:**

Figure 4 - WEBSITE dataset example vulnerable

Scenario 2:

A website has been detected and probably infected with an exploit kit. After a previous analysis cannot be decided if it's malicious. So this uri is sending as a report to the CCH (with a confidence_level = 0.5) in order to feed and to leave this work for analyzers.

Analyzers should take this reports in order to update them. It can be a malicious or a clean uri, in both cases they have to update the confidence_level (malicious > 0.5, clean < 0.5) field.

Dataset sent for scenario 2:

The dataset that take place on this scenario are: **Suspicious and malicious:**

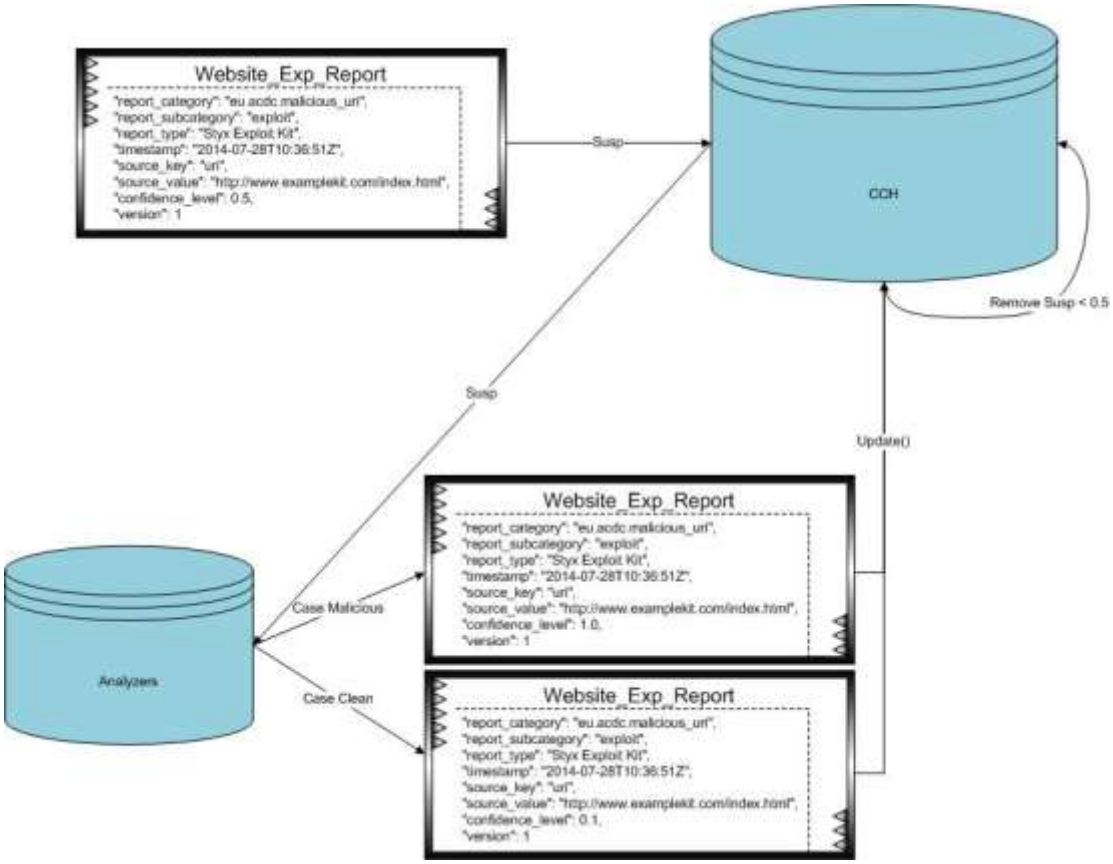

Figure 5 - WEBSITE dataset example suspicious

4.6. Metrics

Depending of the role of each participant the following metrics must be reported in the intermediate reports through the template defined by experiment leaders:

Experiment Phase	Metric	Description	Classified by (if applies)
General	Partners	Number of partners participating on the experiment	• Type of organization • Role in the experiment • Technologies
Deployment & Integration	Tools	Number of tools contributing to the experiment	• Number of deployments • Contribution type
Detection & analysis	WEBSITE volume	Number of websites detected and analyzed	• Detection Tool • ASN • Country • Suspicious • Malicious and Malicious subcategories • Vulnerable
	Website bots	Number of bots attacking websites identified (IP+TS)	• Detection Tool • ASN • Country
	C&C	Number of C&C servers identified on the experiment scope	• Detection Tool • ASN • Country
	MW in websites	MW distributed from websites	• Total • Total analysed • Total Malicious
	Botnets	Number of different botnets detected	• Total
Data Storage	Total WEBSITE Reports in CCH	Number of reports sent to CCH related to websites	• Total accumulated • Per day/week • Per Tool/Partner
Distribution for notification & mitigation purposes	WEBSITE Reports retrieved	Number of website reports retrieved for analysis, notification and mitigation	• Total per partner ASNs (depending the ISP, network owner or CERT constituency) • Per type of element retrieved • For ISPs: Classification per type of network affected (mobile or fixed)
Notification	Notifications	Notifications sent to affected parties. Processes activated with LEAs.	• Total • ASN • Type of element • Sent to: end-user, webmaster, ISP, LEA.

Mitigation	WEBSITE prevention and mitigation contents/tools	Number of visits/downloads to website contents/tools in NSCs	• NSC • Month

Table 34 - WEBSITE metrics

4.7. Reports

During the execution period of the experiment, each participant must complete and send a periodic report (PR) to experiment leaders.

Depending the role in the experiment and the tools operated, this report must contain:

- The metrics
- Incidents or problems during the period
- Specific considerations and conclusions

By default, the PR will be sent weekly, unless a different periodicity could be needed.

Experiment leaders will send a Periodic Report Template per experiment to each participant.

The report Template: **ACDC_EXP_WEBSITE_PR_template.xls** (annexed to this document) will be available also through the Community Portal website.

A final and global report will be developed by experiment leaders. Main conclusions and results will be published on the CP website at the end of each experiment.

5. FAST-FLUX experiment design

The design of this experiment is defined to achieve objectives detailed in section 5.1 of document D3.1-Planning of Experiments.

5.1. Specific definitions for FASTFLUX experiment

Taking in mind that this experiment is focused on detection an analysis of domains using fastflux techniques to support botnet activities, the following terms are defined in the scope of the experiment:

FASTFLUX DOMAIN	
DEFINITION	DNS domain configured in such a way that can hide malicious botnet elements (like phishing websites, malware delivery, etc) using the IP address of multiple compromised devices acting as proxies. The basic idea behind Fast flux is to have numerous IP addresses associated with a single fully qualified domain name, where the IP addresses are swapped in and out with extremely high frequency, through changing DNS records. Those IP address correspond in 99% of cases to infected end-user devices (acting as fastflux bots).
DETECTION	Different technologies and criteria can be used to identify fastflux domains in the experiment: - Single flux detection algorithms. - Double flux detection algorithms.
DATASET	Fast-flux domain dataset

Table 35 - FASTFLUX definition - DOMAIN

FASTFLUX BOT	
DEFINITION	The automated program or piece of malware that control and end-user device to act as proxy of some illegal site controlled by a botnet. In this experiment the fastflux bot is identified by at least a public IP address and the TIMESTAMP of the detection of the fastflux activity on a specific fastflux domain. To identify a Bot it is not necessary to observe it actively participating in an attack.
DETECTION	Fastflux bots are identified through the periodic analysis of the fastflux domains DNS records.
DATASET	Fast-flux bot data set

Table 36 - FASTFLUX definition - BOT

C&C SERVER	
DEFINITION	A C&C server controlling domains configured as fastflux or a server controlling the fastflux bots activities.
DETECTION	C&C servers can be detected from the analysis of: - bot reversing analysis - Fastflux domains/nodes monitoring, correlation and analysis
DATASET	Fast-flux C&C dataset

Table 37 - FASTFLUX definition - C&C

5.1.1. Confidence Level of the information

Independently of the type of element or incident identified, each report shared through the Central Clearing House (CCH) must indicate the level of veracity of the information (through the ***confidence_level*** parameter on the datasets). This is very important for the notification and mitigation part of the experiment.

Common criteria can be applied following guidelines in section 10 of this document.

5.2. Experiment processes and activities

5.2.1. Detection and analysis

The following table details the process and activities to execute along the experiment time (process 1 to 4 are the same for all experiment as defined in section 2). **This table covers detection and analysis activities**.
Not all the activities must be performed by the role identified. Inside the experiment each participant defines the scope of its role and therefore the scope of the actions to execute.

Specific Processes		Description	Activities		Role[5]	Input Info	Output Info
FF1	**Tool detection phase: Detection of Domains using Fast-Flux**	Check whether a domain is using fast-flux techniques and identify resources involved.	**FF1.1**	Check if a domain is using Fast-Flux techniques	Tool Operator	Any input data source used by the tool (including suspicious list of domains from the CCH). FASTFLUX datasets Schemas definition: • Fast-Flux domain. • Fast-Flux C&C. • Fast-Flux botnet. • Fast-Flux bot.	Reports with the data obtained (based on dataset schemata defined).
			FF1.2	Obtain the IP's and timestamp involved in the domain using Fast-Flux			
FF2	**Tool detection phase: Classification and analysis of the data**	Classify the domains using Fast-Flux obtained during process FF1 and analyze them to identify another botnet elements if possible.	**FF2.1**	Analyze & classify botnet activity type (if possible)	Tool Operator	Information collected from process FF1.	Reports with the data obtained. Fastflux domain activity type.
FF3	**Data Correlation**	Correlation of data in order to increase Fast-Flux detections, new rules and events	**FF3.1**	Correlate the data detected and shared by all partners	Tool Operator	Data extracted from the processes FF1 & FF2	Reports with the data obtained (based on dataset schemata defined).

[5] Roles are defined in Document D3.1

FF4	**Delivery data to CCH**	Delivery to the CCH all data and information collected in previous phases	**FF4.1**	Send information obtained to the CCH.	Tool Operator	Information collected and correlated (if apply) from processes FF1, FF2 & FF3 (following partner schema defined in process 3)	
FF5	**Periodic Control Report (Detection & Analysis report by tool)**	Generate a periodic report in order to keep track of the experiment with the information obtained during the experiment detection phase It must be sent to the experiment coordinator at stipulated intervals	**FF5.1**	Generate the report with the specific metrics defined following the template supplied by leaders	Tool Operator	Data from processes FF1, FF2, FF3 & FF4. Periodic Report Template (Detection & Analysis phase)	Periodic Control Report (Detection & Analysis report by tool)
			FF5.2	Send the report to the experiment leaders (INTECO & ATOS)			
FF6	**CCH Monthly Report**	Periodically, generate a report with global Fast-Flux metrics	**FF6.1**	Generate a report with metrics containing the information received, analyzed and collected from the CCH during the last month regarding Fast-Flux	CCH Operator	Information in the CCH. Inputs and outputs requests by partners.	CCH Report
			FF6.2	Send the report to the experiment leaders			

Table 38 - FASTFLUX process detection and analysis

5.2.2. Notification and mitigation

Notification and mitigation activities are very similar on design along the different experiments, so these activities are explained for all in section 8 of this document.

5.2.3. Response times

Some activities of the experiment require maximum response times in order the whole process to be effective. This response times are defined for all experiments in section 9 of this document.

5.3. Experiment Data Flow Diagram

The following diagram shows the dataset flow between roles along the different phases or process of the experiment:

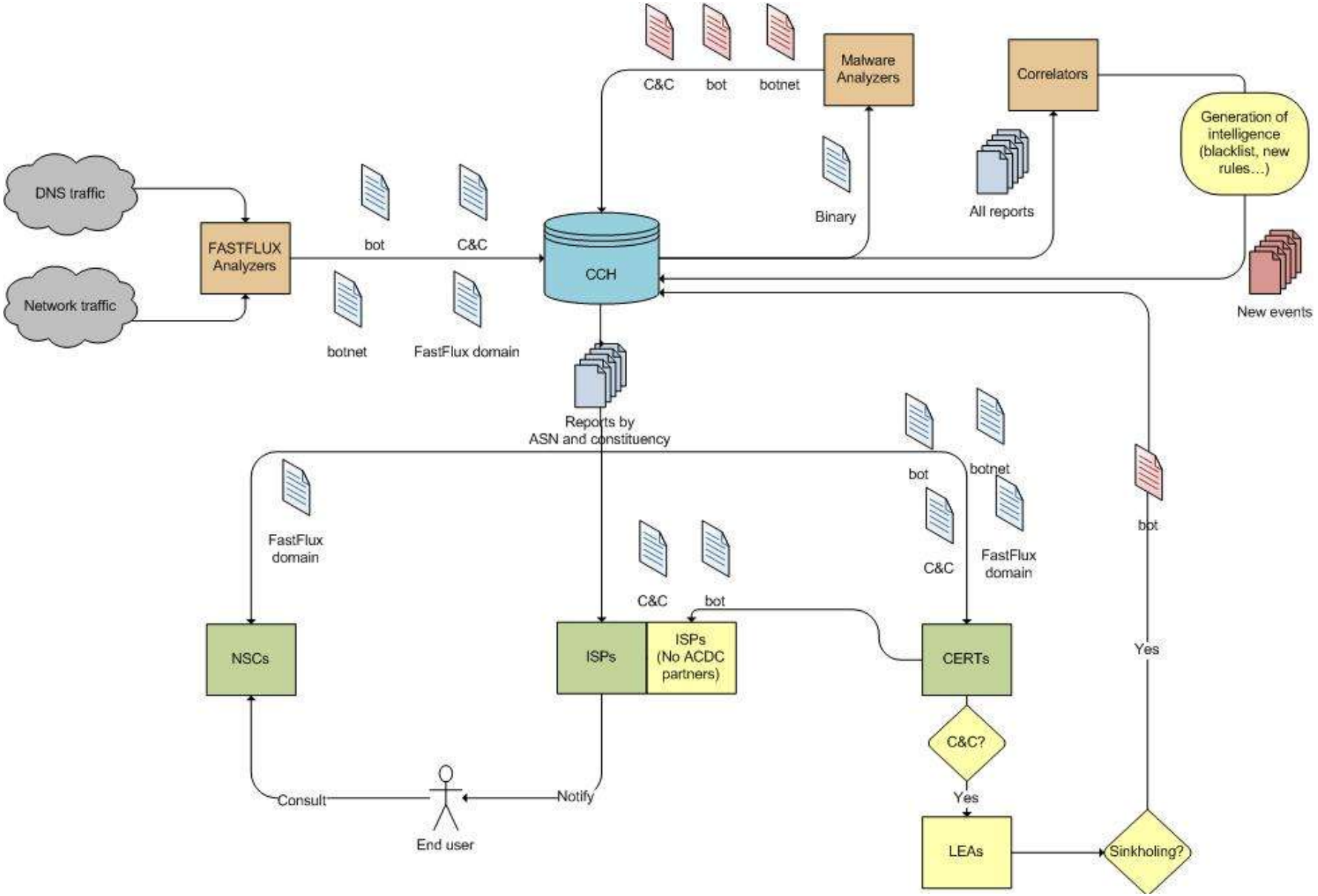

Figure 6 - FASTFLUX experiment data flow

5.4. Datasets definition for FASTFLUX experiment

Based on the specific spam elements to detect and analyse in the scope of the experiment (section 5.1), and on the data schemata defined at the Document D1.7.2 Data Formats Specification, the following datasets has been defined:

- Fast-Flux domain.
- Fast-Flux C&C.
- Fast-Flux botnet.
- Fast-Flux bot.

The fields defined in each dataset are the minimum data for the experiments but they could be extended and any other field can be added by participants.

Extended datasets used must be defined and published through the Community Portal in order to be known by all participants on the experiment.

The following tables contains, for each field defined: a functional description, the field name, the type, and its obligation. In fields with multiple possible values there are specified only those that are involved in this experiment. It also includes some optional fields that are not necessary to send if they are not known.

5.4.1. Fast-flux domain dataset

The following dataset represent the minimum specific data that must be sent for each fast-flux domain.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.fast_flux	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: a domain uri	source_key	string enum: uri	False
The fast flux domain uri	source_value	string format: uri	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	False
The version number of the data format used for the report.	version	integer enum: 1	False

The botnet the fast flux domain can be attributed to.	botnet	string	False
The IP of the associated bot.	src_ip_v4	string format: ipv4	False
	src_ip_v6	string format: ipv6	False

Table 39 - FASTFLUX dataset domain

5.4.2. *Fast-flux C&C dataset*

The following dataset represent the minimum specific data that must be sent for each fast-flux C&C.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.c2_server	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string Format: date-time	False
The type of the reported object	source_key	string enum: ip	False
The IP address of the C&C server	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	False
The version number of the data format used for the report.	version	integer enum: 1	False
The control channel used by the C2.	report_subcategory	string enum: http, irc, other	False
The botnet associated to the C&C.	botnet	string	True

Table 40 - FASTFLUX dataset C&C

5.4.3. *Fast-flux botnet dataset*

The following dataset represent the minimum specific data that must be sent for each spam botnet.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.botnet	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: botnet	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The version number of the data format used for the report.	version	integer enum: 1	True
The category of the botnet	report_subcategory	string enum: c2, p2p, other	True

Table 41 - FASTFLUX dataset botnet

5.4.4. *Fast-flux bot dataset*

The following dataset represent the minimum specific data that must be sent for each bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.bot	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The timestamp when the reported observation took place.	timestamp	string format: date-time	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: IP	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The level of confidence put into the accuracy of the report. A number	confidence_level	number minimum: 0.0	True

between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.		maximum: 1.0	
The version number of the data format used for the report.	version	integer enum: 1	True
The botnet the bot is attributed to.	botnet	string	True
The IP of the C&C where the bot is involved.	c2_ip_v4	string format: ipv4	False
	c2_ip_v6	string format: ipv6	False

Table 42 - FASTFLUX dataset bot

5.5. *Dataset examples*

A functional example of the main dataset flows for this experiment is:

Scenario 1:

A domain "ffdomain.org" has been detected using fast-flux techniques. After tracking the domain a few days, it is discovered that this domain resolve until 5 different IPs:
- 10.10.10.1
- 10.20.10.1
- 10.30.10.1
- 10.40.10.1
- 10.50.10.1

Dataset sent for scenario 1:

The dataset that take place on this scenario is: **domain:**

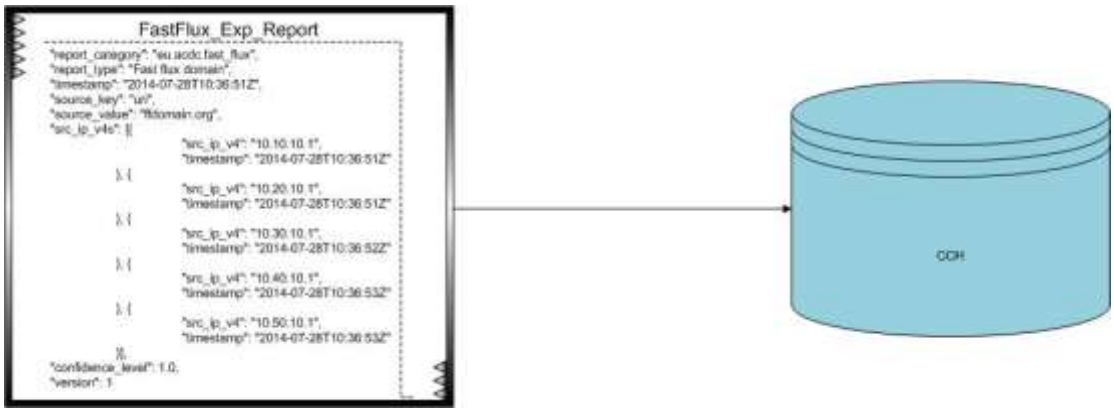

Figure 7 - FASTFLUX dataset example

5.6. Metrics

Depending of the role of each participant the following metrics must be reported in the intermediate reports through the Templates defined by experiment leaders:

Experiment Phase	Metric	Description	Classified by (if applies)
General	Partners	Number of partners participating on the experiment	• Type of organization • Role in the experiment • Technologies
Deployment & Integration	Tools	Number of tools contributing to the experiment	• Number of deployments • Contribution type
Detection & analysis	Fastflux domains	Number of fastflux domains detected and analyzed	• Detection Tool • TLD
	Fastflux bots	Number of total IP addresses used in fastflux techniques	• Detection Tool • Domain • ASN • Country
	C&C	Number of C&C servers identified on the experiment scope	• Detection Tool • ASN • Country
	Botnets	Number of different botnets detected	• Total
Data Storage	Total FASTFLUX Reports in CCH	Number of reports sent to CCH related to FASTFLUX	• Total accumulated • Per day/week • Per Tool/Partner
Distribution for notification & mitigation purposes	Fastflux Reports retrieved	Number of fastflux reports retrieved for analysis, notification and mitigation	• Total per partner ASNs (depending the ISP, network owner or CERT constituency) • Per type of element retrieved (domain, bots) • For ISPs: Classification per type of network affected (mobile or fixed)
Notification	Notifications	Notifications sent to affected parties and processes activated with LEAs.	• Total • ASN • Type of element (domain, bot, C&C) • Sent to: end-user, domain registrar, ISP, LEA.
Mitigation	Fastflux prevention and mitigation	Number of visits/downloads to Fastflux related	• NSC • Month

	contents/tools	contents/tools in NSCs	

Figure 8 - FASTFLUX metrics

5.7. Reports

During the execution period of the experiment, each participant must complete and send a periodic report (PR) to experiment leaders.

Depending the role in the experiment and the tools operated, this report must contain:

- The metrics
- Incidents or problems during the period
- Specific considerations and conclusions

By default, the PR will be sent weekly, unless a different periodicity could be needed.

Experiment leaders will send a Periodic Report Template per experiment to each participant.

The report Template: **ACDC_EXP_FASTFLUX_PR_template.xls** (annexed to this document) will be available also through the Community Portal website.

A final and global report will be developed by experiment leaders. Main conclusions and results will be published on the CP website at the end of each experiment.

6. DDoS experiment design

The design of this experiment is defined to achieve objectives detailed in section 6.1 of document D3.1-Planning of Experiments.

6.1. Specific definitions for DDOS experiment

Taking in mind that this experiment is focused on detection and mitigation of DDOS botnets, the following terms are defined in the scope of the experiment:

DDOS ATTACK	
DEFINITION	A host discovered doing a DDoS attack to another one. It's not an ongoing attack it could be discovered after the attack. In this experiment a DDoS attack is identified by at least a public IP address of the attacker and the TIMESTAMP of the detection of the DDoS attack activity.
DETECTION	Detection of DDoS bots can be done: - Processing and analyzing DDoS Logs from real targets. - Network behaviour analysis - DNS traffic analysis - Identifying attacks to honeynets - ..
DATASET	DDoS attack dataset

Table 43 - DDoS definition - ATTACK

DDOS BOT	
DEFINITION	The automated program or piece of malware installed on an end-user device that is performing automated requests to an internet service (target), being part of a DDoS attack. In this experiment a DDoS bot is identified by at least a public IP address and the TIMESTAMP of the detection of the DDoS attack activity. To identify a Bot it is not necessary to observe it actively participating in an attack.
DETECTION	Detection of DDoS bots can be done: - Sinkholing. - Network behaviour analysis - DNS traffic analysis - ..
DATASET	DDoS bot dataset

Table 44 - DDoS definition - BOT

DDOS C&C SERVER	
DEFINITION	A C&C server of a botnet focused on DDoS attacks.
DETECTION	C&C servers can be detected from the analysis of: - Execution of DDoS bots in dynamic analysis environment.
DATASET	DDoS C&C dataset

Table 45 - DDoS definition - C&C SERVER

6.1.1. Confidence Level of the information

Independently of the type of element or incident identified, each report shared through the Central Clearing House (CCH) must indicate the level of veracity of the information (through the ***confidence_level*** parameter on the datasets). This is very important for the notification and mitigation part of the experiment.

Common criteria can be applied following guidelines in section 10 of this document.

6.2. Experiment processes and activities

6.2.1. Detection and analysis

The following table details the process and activities to execute along the experiment time (process 1 to 4 are the same for all experiment as defined in section 2). **This table covers detection and analysis activities**.
Not all the activities must be performed by the role identified. Inside the experiment each participant defines the scope of its role and therefore the scope of the actions to execute.

Specific Processes		Description	Activities		Role[6]	Input Info	Output Info
DD1	**Tool detection phase: Collecting DDoS attacks data**	Collect RAW data of DDoS attacks.	**DD1.1**	Passive detection: log files	Tool Operator	Log files from target machines or from honeynets. Network or DNS traffic. DDOS datasets Schemas definition: • DDoS attack • DDoS C&C. • DDoS botnet. • DDoS bot.	Reports with DDoS attack raw data.
			DD1.2	Active detection: network traffic			
DD2	**Tool detection phase: Classification, analysis and identification of DDoS-botnets related elements**	The DDoS attacks information and traffic collected in the previous process must be analyzed to identify and classify DDoS-botnet elements	**DD2.1**	Identification of DDoS bots from the analysis of the attack RAW data.	Tool Operators	Information collected from process DD1. DDOS datasets Schemas definition: • DDoS attack • DDoS C&C. • DDoS botnet. • DDoS bot.	Reports with the data obtained (based on dataset schemata defined).
			DD2.2	Analysis of attack payload and/or bot samples to identify C&C servers if possible.			

[6] Roles are defined in Document D3.1

DD3	**Data Correlation**	Correlation of data in order to increase DDoS detections and new rules and events	**DD3.1**	Correlate the data detected and shared by all partners	Tool Operator	Data extracted from processes DD1 & DD2.	Reports with the data obtained (based on dataset schemata defined).
DD4	**Delivery data to CCH**	Delivery to the CCH all data and information collected in previous phases	**DD4.1**	Send information obtained to the CCH.	Tool Operator	Information collected and correlated (if apply) from processes DD1, DD2, DD3. Reports with the data obtained (based on dataset schemata defined).	
DD5	**Periodic Control Report (Detection & Analysis report by tool)**	Generate a periodic report in order to keep track of the experiment with the information obtained during the experiment detection phase It must be sent to the experiment coordinator with the frequency stipulated	**DD5.1**	Generate the report with the specific metrics define in processes DD1, DD2, DD3, and DD4 following the template supplied by leaders	Tool Operator	Data from processes DD1, DD2, DD3, DD4 Periodic Report Template (Detection & Analysis phase)	Periodic Control Report (Detection & Analysis report by tool)
			DD5.2	Send the report to the experiment leaders (INTECO & DE-CIX)			
DD6	**CCH Monthly Report**	Periodically, generate a report with global DDoS metrics	**DD6.1**	Generate a report with metrics containing the information received, analyzed and collected in the CCH during the last month regarding DDoS.	CCH Operator	Information in the CCH. Inputs and outputs requests by partners.	CCH Report
			DD6.2	Send the report to the experiment leaders			

Table 46 – DDoS Process detection and analysis

6.2.2. Notification and mitigation

Notification and mitigation activities are very similar on design along the different experiments, so these activities are explained for all in section 8 of this document.

6.2.3. Response times

Some activities of the experiment require maximum response times in order the whole process to be effective. This response times are defined for all experiments in section 9 of this document.

6.3. Experiment Data Flow Diagram

The following diagram shows the dataset flow between different components along the different phases or process of the experiment:

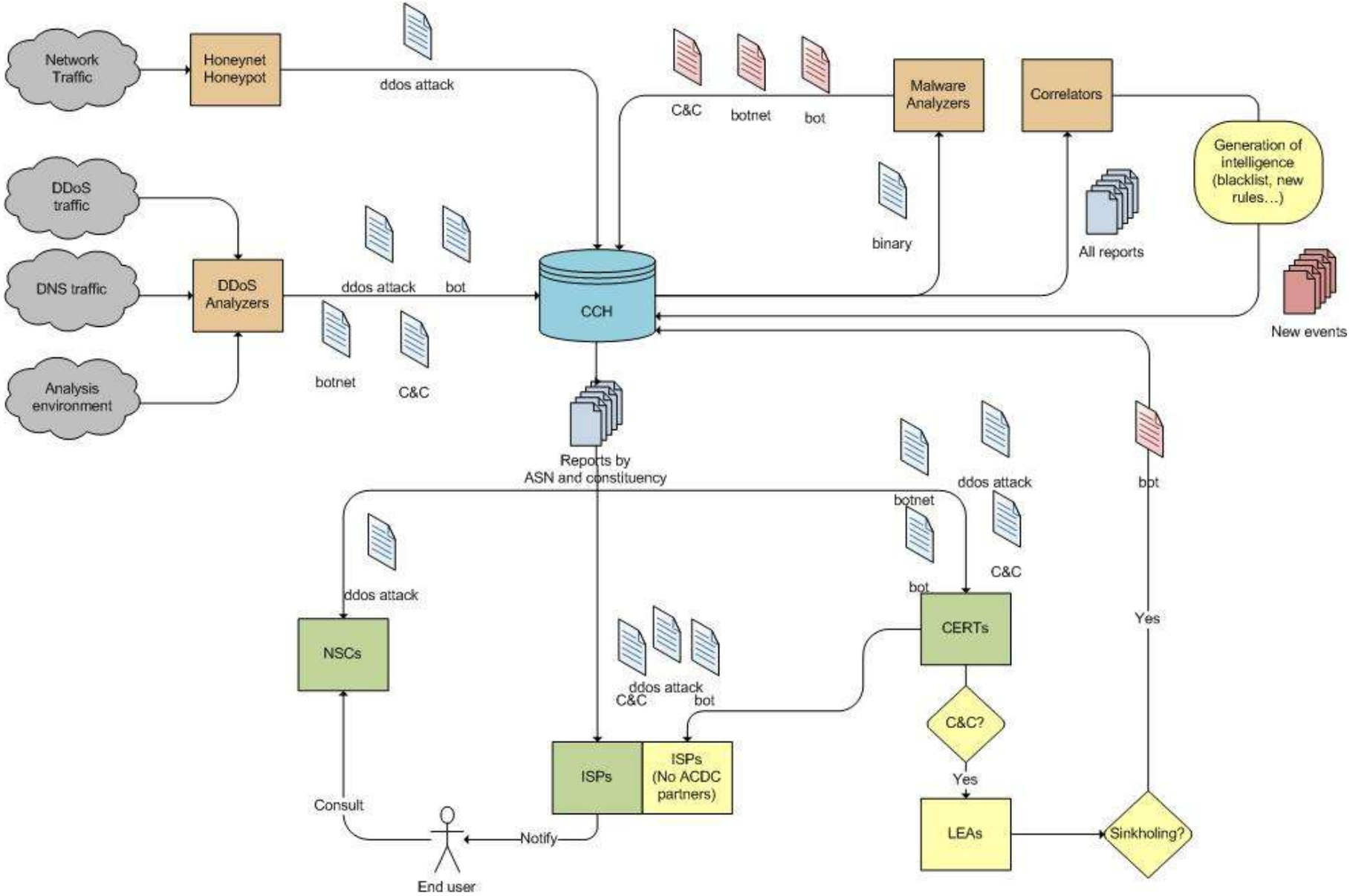

Figure 9 - DDoS experiment data flow

6.4. Datasets for DDOS experiment

Based on the specific spam elements to detect and analyse in the scope of the experiment (section 6.1), and on the data schemata defined at the Document D1.7.2 Data Formats Specification, the following datasets has been defined:

- DDoS attack
- DDoS C&C.
- DDoS bot.
- DDoS botnet.

The fields defined in each dataset are the minimum data for the experiments but they could be extended and any other field can be added by participants.

Extended datasets used must be defined and published through the Community Portal in order to be known by all participants on the experiment.

The following tables contains, for each field defined: a functional description, the field name, the type, and its obligation. In fields with multiple possible values there are specified only those that are involved in this experiment. It also includes some optional fields that are not necessary to send if they are not known.

6.4.1. DDoS attack dataset

The following dataset represent the minimum specific data that must be sent for each DDoS attack.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.attack	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object	source_key	string enum: ip	False
IP of the system performing the attack.	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being	confidence_level	number minimum: 0.0 maximum: 1.0	False

verified to be accurate.			
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the attack performed.	report_subcategory	String enum: dos, dos.dns, dos.http, dos.tcp, dos.udp	False
The RFC 790 decimal internet protocol number of the attack connection.	ip_protocol_number	integer minimum: 0 maximum: 255	False
The IP version of the attack connection.	ip_version	integer enum: 4,6	False
The botnet the attack can be attributed to	botnet	string	True
The source IP of the attack connection. This is always the IP of the attacking system. This field equals source_value.	src_ip_v4	string format: ipv4	True
	src_ip_v6	string format: ipv6	True
The destination port of the attack connection.	dst_port	integer	False

Table 47 - DDoS dataset attack

6.4.2. DDoS C&C dataset

The following dataset represent the minimum specific data that must be sent for each DDoS C&C.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.c2_server	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string Format: date-time	False
The type of the reported object	source_key	string enum: ip	False
The IP address of the C&C server	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being	confidence_level	number minimum: 0.0 maximum: 1.0	False

unreliable and 1.0 being verified to be accurate.			
The version number of the data format used for the report.	version	integer enum: 1	False
The control channel used by the C2.	report_subcategory	string enum: http, irc, other	False
The botnet associated to the C&C.	botnet	string	True

Table 48 - DDoS dataset C&C

6.4.3. DDoS botnet dataset

The following dataset represent the minimum specific data that must be sent for each DDoS botnet.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.botnet	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: botnet	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The version number of the data format used for the report.	version	integer enum: 1	True
The category of the botnet	report_subcategory	string enum: c2, p2p, other	True

Table 49 - DDoS dataset botnet

6.4.4. DDoS bot dataset

The following dataset represent the minimum specific data that must be sent for each DDoS bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.bot	True
The type of the report. This is a free text field characterising the report that should be used for a human	report_type	string	True

readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.			
The timestamp when the reported observation took place.	timestamp	string format: date-time	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: IP	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	True
The version number of the data format used for the report.	version	integer enum: 1	True
The botnet the bot is attributed to.	botnet	string	True
The IP of the C&C where the bot is involved.	c2_ip_v4	string format: ipv4	False
	c2_ip_v6	string format: ipv6	False

Table 50 - DDoS dataset bot

6.5. Dataset examples

A functional example of the main dataset flows for this experiment is:

Scenario 1:

After a DDoS attack it's analyzed in order to separate malicious traffic for benign traffic. Are got the following malicious IPs involved in the attack:
- 10.10.10.1
- 10.10.20.1
- 10.10.30.1
- 10.10.40.1
- 10.10.50.1

Dataset sent for scenario 1:

The dataset that take place on this scenario is: **attack:**

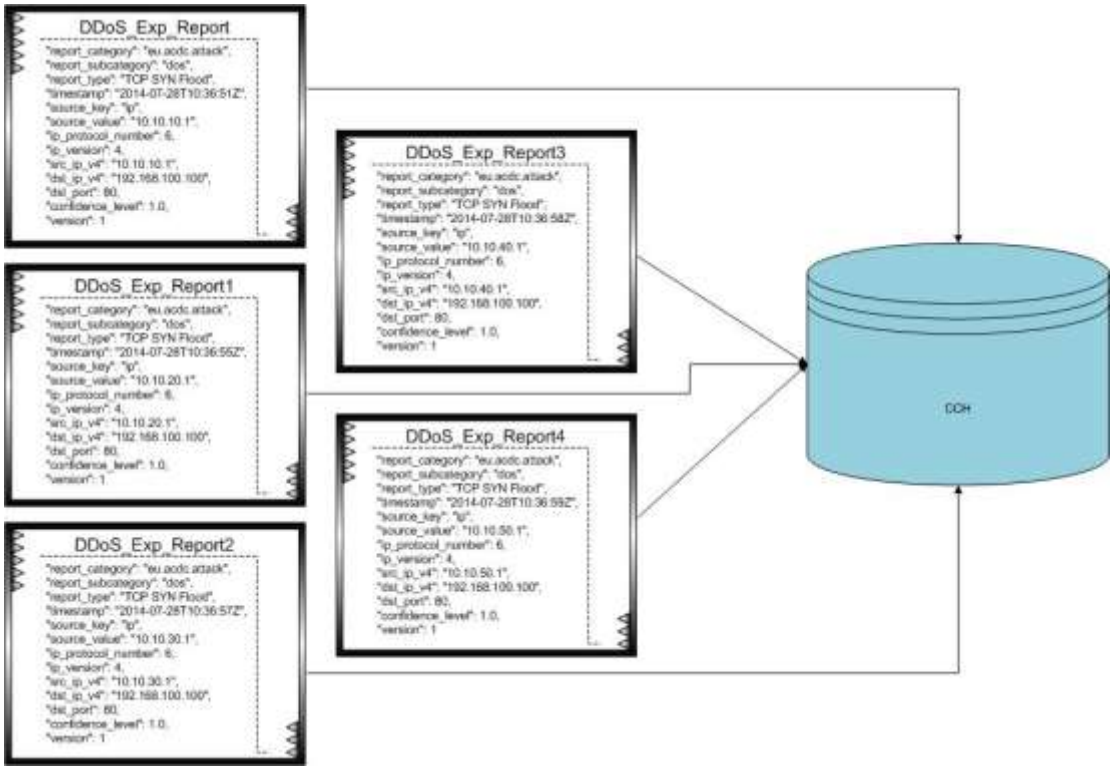

Figure 10 - DDoS dataset example

6.6. Metrics

Depending of the role of each participant the following metrics must be reported in the intermediate reports through the Templates defined by experiment leaders:

Experiment Phase	Metric	Description	Classified by (if applies)
General	Partners	Number of partners participating on the experiment	• Type of organization • Role in the experiment • Technologies
Deployment & Integration	Tools	Number of tools contributing to the experiment	• Number of deployments • Contribution type
Detection & analysis	DDoS attacks	Number of DDoS real attacks analyzed	• Input source, partner or service who detects the attack.
	DDoS bots	Number of total IP addresses identified as DDoS bots	• Attack (identifier) • ASN • Country
	C&C	Number of C&C servers identified on the experiment scope	• Detection Tool • ASN • Country
	Botnets	Number of different botnets detected	• Total
Data Storage	Total DDoS Reports in CCH	Number of reports sent to CCH related to	• Total accumulated • Per day/week

		DDoS	• Per Tool/Partner
Distribution for notification & mitigation purposes	DDoS Reports retrieved	Number of DDoS reports retrieved for analysis, notification and mitigation	• Total per partner ASNs (depending the ISP, network owner or CERT constituency) • Per type of element retrieved (C&C, bots) • For ISPs: Classification per type of network affected (mobile or fixed)
Notification	Notifications	Notifications sent to affected parties and processes activated with LEAs.	• Total • ASN • Type of element (bot, C&C) • Sent to: end-user, ISP, LEA.
Mitigation	DDoS prevention and mitigation contents/services	Number of visits to DDoS related contents/services in NSCs	• NSC • Month

Table 51 - DDoS metrics

6.7. Reports

During the execution period of the experiment, each participant must complete and send a periodic report (PR) to experiment leaders.

Depending the role in the experiment and the tools operated, this report must contain:

- The metrics
- Incidents or problems during the period
- Specific considerations and conclusions

By default, the PR will be sent weekly, unless a different periodicity could be needed.

Experiment leaders will send a Periodic Report Template per experiment to each participant.

The report Template: **ACDC_EXP_DDOS_PR_template.xls** (annexed to this document) will be available also through the Community Portal website.

A final and global report will be developed by experiment leaders. Main conclusions and results will be published on the CP website at the end of each experiment.

7. MOBILE experiment design

The design of this experiment is defined to achieve objectives detailed in section 7.1 of document D3.1-Planning of Experiments.

7.1. Specific definitions for MOBILE experiment

Taking in mind that this experiment is focused on detection an analysis of security incidents on mobile devices, the following terms are defined in the scope of the experiment:

MOBILE SUSPICIOUS	
DEFINITION	APKs found on the device and performing suspicious[7] activities or suspicious[8] urls used from the device. These elements must be further analysed to discern if they are malicious or not.
DETECTION	Both types of elements within this category should be analyzed by a URL analyzer or a malware analyzer in the scope of the experiment.
DATASETS	Mobile suspicious elements dataset

Table 52 - MOBILE definition - SUSPICIOUS

MOBILE MALICIOUS	
DEFINITION	APKs found on the device and performing malicious activities (for instance, credential thief) or malicious urls used from the device.
DETECTION	URL or malware analysis must be done to report the elements as malicious.
DATASETS	Mobile malicious elements dataset

Table 53 - MOBILE definition - MALICIOUS

MOBILE BOT	
DEFINITION	Mobile device compromised or infected with malware and controlled by a botnet to perform specific illegal activities, such as malware distribution or monetize illegal activities for instance sending SMS premium. In the experiment a mobile bot is identified by at least a public IP address and the TIMESTAMP of the detection of the malicious activity. To identify a Bot it is not necessary to observe it actively participating in an attack.
DETECTION	Different technologies and criteria can be used to identify mobile bots in the experiment: - IDSs checking outgoing connections done from the mobile device. - Use of automated SMS premium or calls to premium services. - APK analysis. - Sinkholing - Etc.

[7] Is consider as suspicious if there are enough evidences of it is doing any malicious activity but it needs a deeper analysis to confirm it.

[8] Is consider as suspicious if there are enough evidences of it is doing any malicious activity but it needs a deeper analysis to confirm it.

MOBILE BOT	
DATASET	Mobile bot dataset

Table 54 - MOBILE definition - BOT

MOBILE ATTACK	
DEFINITION	Mobile device doing illegal activities, such as malware distribution or monetize illegal activities for instance sending SMS premium. In the experiment a mobile attack is identified by at least a public IP address of the system performing the attack and the TIMESTAMP of the detection of the malicious activity.
DETECTION	Different technologies and criteria can be used to identify mobile bots in the experiment: - IDSs checking outgoing connections done from the mobile device. - Use of automated SMS premium or calls to premium services. - APK analysis. - Etc.
DATASET	Mobile attack dataset

Table 55 - MOBILE definition - ATTACK

C&C SERVER	
DEFINITION	A C&C server member of a botnet focused on control infected mobile devices and malicious mobile activities (like malware distribution, fraud or any illegal activity). In the scope of the experiment, also a C&C server found as a result of the analysis of the different elements detected on the mobile experiment.
DETECTION	C&C servers can be detected from the analysis of: - Malicious APKs found in the mobile devices - Mobile bot reversing analysis - Analysis of the outgoing connections done from the mobile devices - Correlation activities.
DATASET	Mobile C&C dataset

Table 56 - MOBILE definition - C&C

7.1.1. Confidence Level of the information

Independently of the type of element or incident identified, each report shared through the Central Clearing House (CCH) must indicate the level of veracity of the information (through the ***confidence_level*** parameter on the datasets). This is very important for the notification and mitigation part of the experiment.

Common criteria can be applied following guidelines in section 10 of this document.

7.2. Experiment process and activities

7.2.1. Detection and analysis

The following table details the process and activities to execute along the experiment time (process 1 to 4 are the same for all experiment as defined in section 2). **This table covers detection and analysis activities**.
Not all the activities must be performed by the role identified. Inside the experiment each participant defines the scope of its role and therefore the scope of the actions to execute.

Process		Description	Activities			Role[9]	Input Info	Output Info
MB1	**Tool detection phase: Collecting data from CCH**	OPTIONAL: Collect information from the CCH useful to feed systems spam-botnet sensors in order to increase number and quality detections. This process is a constant task through the experiment.	**MB1.1**	Request the necessary information needed based on the datasets available on the CCH		Tool Operator	Datasets available in CCH.	New detection rules for sensors.
			MB1.2	Feed the detection tool with the information obtained				
MB2	**Tool detection phase: data collection**	Data collection from mobile devices (end-user tools and/or network sensors). This data will be used to detect and identify mobile bots and malware specific for mobile, and to obtain valuable data for statistics.	**MB2.1**	Suspicious data collection from traffic network	Attack traffic to device. (IPs, domains, Payload).	Tool Operator	Device information, network generated traffic and APKs installed. MOBILE datasets Schemas definition: • Mobile attack • Mobile suspicious elements • Mobile botnet. • Mobile bot.	IPs of attackers to mobile devices. Payloads. Suspicious APKs
					Malicious SMS			
					Suspicious outgoing connections (to blacklists)			
			MB2.2	Detection of suspicious APKs from mobile devices.				

[9] Roles are defined in Document D3.1

MB3	**Tool detection phase: Analysis & Classification Mobile Botnets**	The Mobile information collected in the previous process must be analyzed to identify and classify Mobile botnet elements	**MB3.1**	Identify botnet elements (C&C and Bots)	Tool Operator	Information collected from process MB2. MOBILE datasets Schemas definition: • Mobile malicious elements. • Mobile C&C. • Mobile botnet.	Reports with the data obtained (based on dataset schemata defined).
			MB3.2	Analysis of APKs. Identify malicious APKs			
			MB3.3	Analysis of outgoing connections			
MB4	**Data correlation**	Correlation of data in order to increase Mobile botnet detections and new rules and events	**MB4.1**	Correlate the data detected and shared by all partners	Tool Operator	Data extracted from processes MB2 & MB3 or other experiments.	Reports with the data obtained (based on dataset schemata defined).
MB5	**Delivery data to CCH**	Delivery to the CCH all data and information collected in previous phases Must be generated the feeds to send to the CCH based on the schemas defined by each partner.	**MB5.1**	Send information obtained to the CCH.	Tool Operator	Information collected and correlated (if apply) from processes MB2, MB3 and MB4 (based on dataset schemata defined).	
MB6	**Periodic Control Report (Detection & Analysis report by tool)**	Generate a periodic report in order to keep track of the experiment with the information obtained during the experiment detection phase It must be sent to the experiment coordinator at stipulated intervals	**MB6.1**	Generate the report with the specific metrics defined, following the template supplied by leaders	Tool Operator	Data from processes MB1, MB2, MB3, MB4 & MB5 Periodic Report Template (Detection & Analysis phase)	Periodic Control Report (Detection & Analysis report by tool)
			MB6.2	Send the report to the experiment leaders (INTECO & XLAB)			

MB7	**CCH Monthly Report**	Periodically, generate a report with global mobile botnet & information metrics	**MB7.1**	Generate a report with metrics containing the information received, analyzed and collected from the CCH during the last month regarding Mobile botnets	CCH Operator	Information in the CCH. Inputs and outputs requests by partners.	CCH Report
			MB7.2	Send report to the experiment leaders			
			MB7.3	Send report to ACDC partners			

Table 57 - MOBILE process detection and analysis

7.2.2. Notification and mitigation

Notification and mitigation activities are very similar on design along the different experiments, so these activities are explained for all in section 8 of this document.

Specific to MOBILE experiment is the analysis and detection of malicious APKs. In this sense, NSCs participants on the experiment should:

- Retrieve malicious APKs from CCH
- Analyze which ones are affecting to users of its country.
- In case positive, generate the content and advertise about it through the NSC web portal (See success criteria defined in D3.1).

An example of this can be found in Annex I.

7.2.3. Response times

Some activities of the experiment require maximum response times in order the whole process to be effective. This response times are defined for all experiments in section 9 of this document.

7.3. Experiment Data Flow Diagram

The following diagram shows the dataset flow between different components along the different phases or process of the experiment:

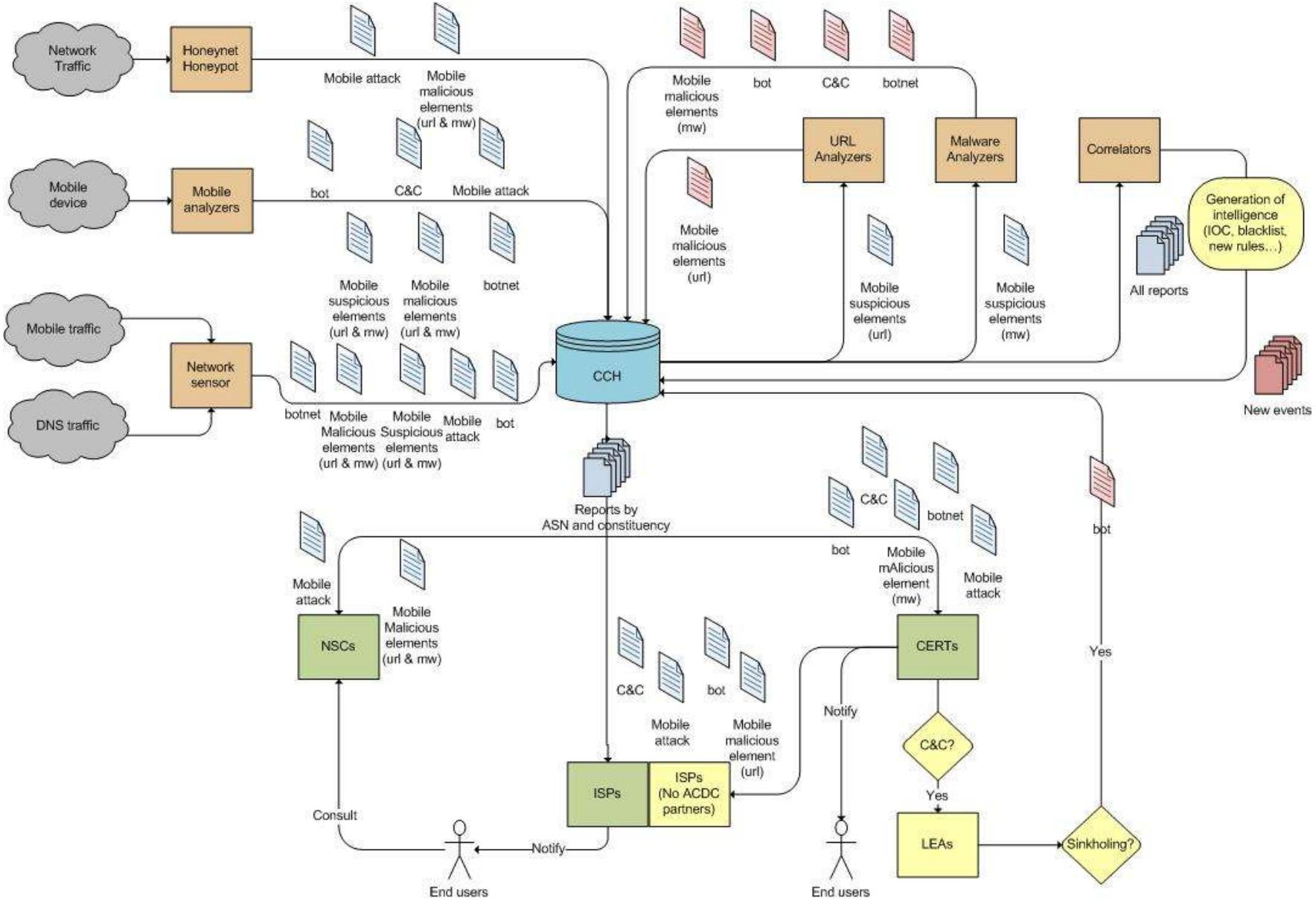

Figure 11 - MOBILE experiment data flow

7.4. Datasets definition for MOBILE experiment

Based on the specific spam elements to detect and analyse in the scope of the experiment (section 7.1), and on the data schemata defined at the Document D1.7.2 Data Formats Specification, the following datasets has been defined:

- Mobile attack.
- Mobile C&C.
- Mobile suspicious elements
- Mobile malicious elements.
- Mobile botnet.
- Mobile bot.

The fields defined in each dataset are the minimum data for the experiments but they could be extended and any other field can be added by participants.

Extended datasets used must be defined and published through the Community Portal in order to be known by all participants on the experiment.

The following tables contains, for each field defined: a functional description, the field name, the type, and its obligation. In fields with multiple possible values there are specified only those that are involved in this experiment. It also includes some optional fields that are not necessary to send if they are not known.

7.4.1. Mobile attack dataset

The following dataset represent the minimum specific data that must be sent for each mobile attack.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.attack	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object.	source_key	string enum: ip	False
IP of the bot.	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	False

The version number of the data format used for the report.	version	integer enum: 1	False
The type of the attack.	report_subcategory	enum: abuse, compromise, data, login, malware, other	False
The RFC 790 decimal internet protocol number of the attack connection.	ip_protocol_number	integer minimum: 0 maximum: 255	False
The IP version of the attack connection.	ip_version	integer enum: 4, 6	False
The botnet the attack can be attributed to (if apply).	botnet	string	True
Additional data for the observation, as the events related with the report.	additional_data	object	True
The IP of the spam bot.	src_ip_v4	string format: ipv4	False
	src_ip_v6	string format: ipv6	False
The destination port of the attack connection.	dst_port	integer	False
The filename used for the payload that the attack tried to install or run on the attacked system.	sample_filename	string	True
The SHA256 hash of the payload that the attack tried to install or run on the attacked system.	sample_sha256	string	True
The URI of the payload in the wild that the attack tried to install or run on the attacked system.	malicious_uri	string format: uri	True

Table 58 - MOBILE dataset attack

7.4.2. Mobile C&C dataset

The following dataset represent the minimum specific data that must be sent for each mobile C&C.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.c2_server	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the	timestamp	string	False

reported observation took place.		Format: date-time	
The type of the reported object	source_key	string enum: ip	False
The IP address of the C&C server	source_value	string	False
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	False
The version number of the data format used for the report.	version	integer enum: 1	False
The control channel used by the C2.	report_subcategory	string enum: http, irc, other	False
The botnet associated to the C&C.	botnet	string	True

Table 59 - MOBILE dataset C&C

7.4.3. Mobile suspicious elements dataset

The following dataset represent the minimum specific data that must be sent for each mobile suspicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

7.4.3.1. Suspicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False

The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious content at the uri.	report_subcategory	string enum: exploit, malware, phishing, other	False
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 60 - MOBILE dataset suspicious uri

7.4.3.2. Suspicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a suspicious 0.5.	confidence_level	number enum: 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True

The binary of the sample encoded in base 64.	sample_b64	string	True

Table 61 - MOBILE dataset suspicious malware

7.4.4. Mobile malicious elements dataset

The following dataset represent the minimum specific data that must be sent for each mobile malicious element. This is composed by two specific data schemata: eu.acdc.malicious_uri and eu.acdc.malware.

7.4.4.1. Malicious URI dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malicious_uri	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: uri	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The type of the malicious content at the uri.	report_subcategory	string enum: exploit, malware, phishing, other.	False
The botnet the malicious uri can be attributed to.	botnet	string	True
For the malicious uri, the file name of the malicious content.	sample_filename	string	True
For the malicious uri, the SHA256 hash of the malicious content.	sample_sha256	string	True

Table 62 - MOBILE dataset malicious uri

7.4.4.2. Malicious malware dataset

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.malware	False
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	False
The timestamp when the reported observation took place.	timestamp	string format: date-time	False
The type of the reported object: an URI or a malware sample.	source_key	string enum: malware	False
The uri to the malicious content or the SHA256 has of the malware sample.	source_value	string	False
The level of confidence put into the accuracy of the report. As a malicious > 0.5.	confidence_level	number enum: > 0.5	False
The version number of the data format used for the report.	version	integer enum: 1	False
The botnet the sample is attributed to.	botnet	string	True
The binary of the sample encoded in base 64.	sample_b64	string	True

Table 63 - MOBILE dataset malicious malware

7.4.5. Mobile botnet dataset

The following dataset represent the minimum specific data that must be sent for each mobile botnet.

Description	Field name	Type	Description
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.botnet	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one	report_type	string	True

sentence.			
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: botnet	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The version number of the data format used for the report.	version	integer enum: 1	True
The category of the botnet	report_subcategory	string enum: c2, p2p, other	True

Table 64 - MOBILE dataset botnet

7.4.6. Mobile bot dataset

The following dataset represent the minimum specific data that must be sent for each mobile bot.

Description	Field name	Type	Optional
The category of the report. This links the report to one of ACDC's schemata.	report_category	string enum: eu.acdc.bot	True
The type of the report. This is a free text field characterising the report that should be used for a human readable description rather than for automatic processing. As a rule of thumb this should not be longer than one sentence.	report_type	string	True
The timestamp when the reported observation took place.	timestamp	string format: date-time	True
The type of the reported object: a botnet or an IP address of the bot.	source_key	string enum: IP	True
The identifier of the botnet or the IP address of the infected system.	source_value	string	True
The level of confidence put into the accuracy of the report. A number between 0.0 and 1.0 with 0.0 being unreliable and 1.0 being verified to be accurate.	confidence_level	number minimum: 0.0 maximum: 1.0	True
The version number of the data format used for the report.	version	integer enum: 1	True
The botnet the bot is attributed to.	botnet	string	True
Additional data for the observation, as the events related with the report.	additional_data	object	True
The IP of the C&C where the bot is involved.	c2_ip_v4	string format: ipv4	False
	c2_ip_v6	string format: ipv6	False

Table 65 - MOBILE dataset bot

7.5. *Dataset examples*

A functional example of the main dataset flows for this experiment is:

Scenario 1:

By several events from IDS, have been detected several bots doing malicious connections:

- 10.20.10.1
- 10.30.10.1
- 10.40.10.1

After study the traffic generated, was realized than the bots are sending information to the C&C:

- 10.10.10.1

Dataset sent for scenario 1:

The dataset that take place on this scenario are: **C&C, botnet and bot:**

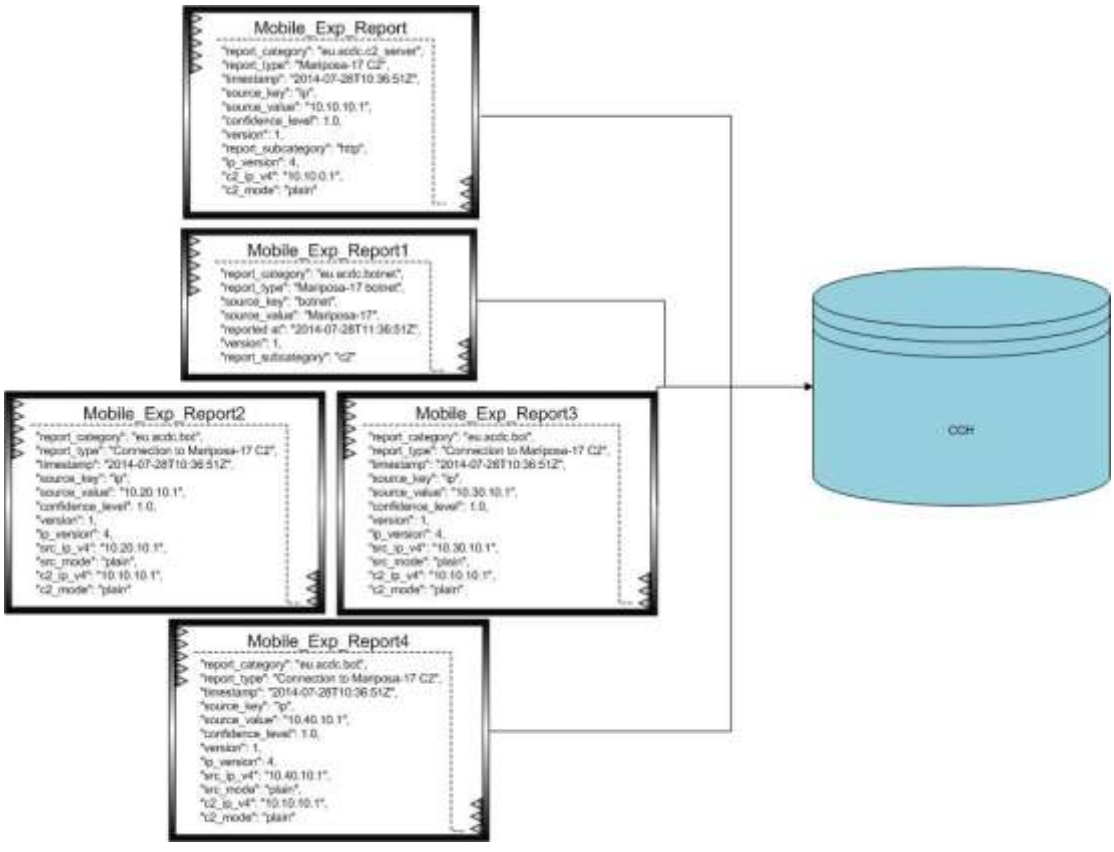

Figure 12 - MOBILE dataset example

7.6. Metrics

Depending of the role of each participant the following metrics must be reported in the intermediate reports through the Templates defined by experiment leaders:

Experiment Phase	Metric	Description	Classified by (if applies)
General	Partners	Number of partners participating on the experiment	• Type of organization • Role in the experiment • Technologies
Deployment & Integration	Tools	Number of tools contributing to the experiment	• Number of deployments • Contribution type
Detection & analysis	Mobile Events	Numbers of mobile events detected	• Detection Tool • Type of event • Suspicious • Malicious (type of activity)
	APKs	Number of APKs analyzed	• Detection Tool • Suspicious • Malicious (type of activity)
	Mobile Bots	Number of mobile bots identified (IP+TS)	• Detection Tool • ASN • Country
	C&C	Number of C&C servers identified on the experiment scope	• Detection Tool • ASN • Country
	Botnets	Number of different botnets detected	• Total
Data Storage	Total MOBILE Reports in CCH	Number of reports sent to CCH related to mobile	• Total accumulated • Per day/week • Per Tool/Partner
Distribution for notification & mitigation purposes	MOBILE Reports retrieved	Number of MOBILE reports retrieved for analysis, notification and mitigation	• Total per partner ASNs (depending the ISP, network owner or CERT constituency) • Per type of element retrieved
Notification	Notifications	Notifications sent to end users and processes activated with LEAs.	• Total • ASN • Type of element • Sent to: end-user, ISP, LEA.
Mitigation	MOBILE prevention and mitigation contents/tools	Number of visits/downloads to MOBILE contents/tools in NSCs	• NSC • Month

Table 66 - MOBILE metrics

7.7. Reports

During the execution period of the experiment, each participant must complete and send a periodic report (PR) to experiment leaders.

Depending the role in the experiment and the tools operated, this report must contain:

- The metrics
- Incidents or problems during the period
- Specific considerations and conclusions

By default, the PR will be sent weekly, unless a different periodicity could be needed.

Experiment leaders will send a Periodic Report Template per experiment to each participant.

The report Template: **ACDC_EXP_MOBILE_PR_template.xls** (annexed to this document) will be available also through the Community Portal website.

A final and global report will be developed by experiment leaders. Main conclusions and results will be published on the CP website at the end of each experiment.

8. Design of mitigation activities (all experiments)

Mitigation activities are very similar on design along the different experiments. These activities are oriented to network owners, ISPs, CERTs and NSCs that are participants of the experiment. Those roles or entities are the ones that can notify or communicate to final affected users, webmasters, by different channels and ways. Also are network owners that can implement prevention, proactive and reactive procedures to mitigate the effects of threats identified along the experiments. National CERTs usually has cybersecurity competencies, so in some cases, can launch official notifications to LEAs, in order to take down malicious servers.

The goal is that periodically (see response times in section 9.1), through the execution time of the experiments, those partners will retrieve relevant information from the CCH that affects its network or constituency: C&Cs, bots, spam campaigns, malicious APKS, etc. With this information analysis, notification, prevention and mitigation activities must be launched and reported. The following table describes the process:

Process		Description	Activities		Role[10]	Input Info	Output Info
M1	**Collect and analyze information from CCH (Analysis by owners of affected resources)**	Network Owners, ISPs and CERTs must retrieve reports related to their network and/or constituency from the CCH in order to analyze them, previously to notification and/or mitigation phase.	M1.1	Collect the corresponding information from the CCH: **Network owners & ISPs**: bots, attacks, C&C and malicious URLs on their ASNs. **CERTs:** Bots and C&C on their constituency. Malware sent by IPs under its constituency. Malicious websites and URLs under its constituency. Vulnerable Websites under their constituency. IOCs or main attacks types if available. **NSCs:** Spam Campaigns.	Network owners, ISPs, CERTs and NSCs.	CCH data	Set of incidents to be notified.

[10] Roles are defined in Document D3.1

				Malicious APKs. Country Bots (if online checking services are in place and if it legally feasible) IOCs or main attacks types if available.			
			M1.2	Process the information, classify it and identify what is going to be notified.			
			M1.3	**ISPs:** Classify the incidents by mobile or fixed networks	ISPs	CCH data and ISP information about network address space.	Periodic Report with the type of incidents affecting mobile networks and fixed networks. Each ISP must define the classification that can be reported.
M2	**Threat Notification**	Notify end users affected by the incidents, motivating them to put a solution to mitigate the threat or be disinfected.	M2.1	Generate the notification according to the threat behavior.	CERTs and ISPs	Information collected in process M1.	Notifications
			M2.2	Send the notification to end users/network or resources owners. The notification must include (if it applies) the address of the National Support Center (depending of each country)	CERTs and ISPs		
			M2.3	In the case of C&Cs identified, activate a notification process with national LEAs in order to take down and control the C&C server to mitigate the botnet.	CERTs (If legally feasible, depending on the CERT's competencies)		
			M2.4	Monitoring notification process.	CERTs and ISPs		
			M2.5	Provide other notification mechanisms like auto-checking online services for end-users through National Support Centers.	NSCs		

M3	**Threat Mitigation**	Create contents, publish disinfection tools or cleaners, advisories or specific services to mitigate or prevent botnet incidents. These contents will be available for end-users through NSC channels (website, social networks, etc.) or other online channels from partners.	M3.1	In general: Publish botnet information contents and specific advisories related to botnet threats and activity. SPAM experiment scope: Publish advisories about **spam campaigns** affecting specific areas or countries. WEBSITE experiment scope: Publish contents related to **main type of attacks to websites**. MOBILE experiment scope: Publish advisories about **malicious APKs for mobile devices**.	NSCs	Data from process M1.	Alerts, advisories or services for mitigation. In case of a sinkholing, for example, new reports with bot data will be sent to CCH to start the cycle of collecting and notification process.
			M3.2	Publish and disseminate cleaners to disinfect botnet malware.	NSCs		
			M3.3	If process with LEAs success, develop mitigation actions against C&C servers like sinkholing or isolating the server for analysis, etc. The results of this action would feed previous processes (for example: bot identification and reporting to CCH).	CERTs		
			M3.4	Implementation of others mitigation actions or services[11]. For example: spammers blocking at network level by ISPs, black lists URL blocking by network owners, etc.	ISPs and Network Owners		

[11] Partners must specify the specific mitigation actions that can be implemented on its networks.

PR_M	**Periodic Control Report (Notification and Mitigation)**	Generate an intermediate report in order to keep track of the experiment with the information obtained during the experiment notification and mitigation phase. It must be sent to the experiment coordinator with the frequency stipulated.	PR_M.1	Generate the report following the template supplied by leaders.	Network owners, ISPs, CERTs and NSCs.	Data from previous process. Periodic Report Template (Notification and Mitigation)	Periodic Control Report (Mitigation)
			PR_M.2	Send the report to the experiment leaders.			

Table 67 - Mitigation activities

Examples of implementation of some mitigation activities can be found on Annex I.

9. Specific Experiments Conditions

In order to achieve the different experiment objectives and success criteria, the following conditions must be met. If any of them cannot be met, the participant must inform the experiment leaders with the reason and possible countermeasures:

9.1. Response Times

- **Report detection activity delay**: Maximum time between botnet elements (bots and C&C) and malicious resources detection time and reporting time to CCH: **one day**.
- **Retrieving periods**:
  - ISPs and CERTs must check at least **every day** for C&C servers, bots, and malicious resources on its networks or constituency.
  - Analyzers must check at least **every day** for suspicious reports that need further analysis to verify the malicious activity.
- **Analysis time**: Maximum time between retrieving suspicious elements from CCH for analysis and sending the results back to CCH: **two days**.
- **Notification times**: Notifications must be done following specific incident management procedures defined by CERT or ISP (those procedures are not in the scope of the project).
- **NSC advisories publication time**: If a specific botnet activity is identified affecting users of a participant NSC, the advisory on the NSC must be published maximum **three days** later from detection.

9.2. Analysis Capacity

Some activities of the experiment require analysis capacities, in some cases analysis can be automated but also manual analysis can be necessary to be done.

Analysis activities involve all partners in different ways. Not all partners has the same resources (human, HW, etc.) assigned to the experiment, so it is difficult in design identify the overall analysis capacity for each experiment, even more when on design, it is not completely close the number of participants on a specific experiment.

Following success criteria defined in D3.1, for example in spam experiment, it is defined the following: ***75% of suspicious files and URLs in spam must be analyzed***.

- ⇨ In general, in order to achieve success criteria for the experiments, all participants must execute the designed activities and report to experiment leaders WHEN & WHY analysis capacities are being exceeded.
- ⇨ Depending on this, experiment leaders can analyze the situation, the impact for the experiment, and define countermeasures.
- ⇨ For example:
  - In some cases, analysis activities could be redistributed along different tools/partners, in order to better distribute the workload during the execution of the experiment.
  - In other cases, high volumes of some kind of information can be reduced if the data is not being contributing in quality to the experiment objectives.
  - In other cases, filters or sampling could be necessary be implemented.

10. Confidence Levels definition

This section describes a common criteria that can be applied or be followed by partners in the scope of the experiment in order to better assign a value to the confidence_level field for each report sent to the CCH.

The following table has been defined as a guideline and follows specifications given in D1.7.2 document of ACDC. The table lists possible criteria to decide if a report is high, medium or low confidence. Level 0 indicates that the report contains NO REAL data, so it is considered experimental.

CONFIDENCE LEVEL	VALUE (From D1.7.2)	GUIDELINES TO CLASSIFY
HIGH/VERIFIED: The sender has enough confidence in the data to classify the information as valid for an analyst and to notify the owner of the reported source.	1.0	✓ The data reported has been analyzed and verified in the project scope. ✓ The reporting source has been accredited as high confidence by the partner. ✓ The data comes from a sinkholing controlled process. ✓ The origin detection tool is high stable and contrasted (very low of false positives)
MEDIUM/SUSPICIOUS: The sender has some indication that the data is correct, like some anomaly detection triggered an alert. The information is not reliable enough for notification but a good source for further analysis.	0.5	✓ The data reported has been analyzed and verified by a third party. ✓ The reporting source has been accredited as medium confidence by the partner. ✓ Data has been obtained from a beta version of a sensor. ✓ Medium level of analysis has been done and reasonably evidence of malicious activity has been found, but the data needs further analysis to be verified and be useful for notification.
LOW/NOT ENOUGH EVIDENCE: The sender has low trust in the data that for example comes from a sensor that is prone to false positives. The information is not reliable enough for notification and should only be used as a second source for confirmation. FALSE POSITIVE: The sender has analyzed previous data reported with medium or high level of confidence, but the analysis confirms a false positive.	0.1	LOW: **THIS DATA MUST NOT BE REPORTED TO CCH** ✓ Data has been obtained from alpha version of a sensor (initial stage of development), so can potentially report a contrasted high level of false positives. FALSE POSITIVE: **THIS DATA MUST BE REMOVED FROM CCH** ✓ Medium or high confidence data previously stored in CCH has been analyzed and classify as false positive by another partner. In this case, the level of confidence must be updated by the analyzers and, once in CCH, this data must be immediately removed[12] by CCH operators.
EXPERIMENTAL: The data is not reliable. The information consists of test data not to be used in any further processing.	0.0	No real data, the data is experimental for the experiment. For example simulated data to help to fine-tune of the detection or reporting tools.

Table 68 – Confidence level definition

[12] Following Legal Requirements.

11. ANNEX I – Mitigation examples

This section contains some examples of notification and mitigation activities that can be a guideline for all experiment participants.

11.1. ISP notification

This is an example of a Spanish ISP informing a client (end-user) that it has been detected spam-bot activities from its internet connection. A reference to INTECO is provided in order to find tools for disinfection:

Estimado cliente:

Desde Centro Némesys le informamos de que estamos recibiendo quejas de otros ususarios comunicándonos la recepción de correos no deseados, procedentes de su línea de acceso a Internet con número XXXXXXXXX.

Por este motivo, como medida de protección, hemos procedido a bloquear el puerto SMTP (25) de envío de correo electrónico.

Las causas más probables de que alguien esté enviando dichos mensajes de correo sin su conocimiento, pueden ser:
- que sus equipos estén contaminados por algún virus o troyano,
- que haya instalado un servidor de correo y no lo tenga bien protegido o configurado, y que a través de él puedan conectarse anónimamente para enviar correos,
- que terceras personas estén utilizando su red WIFI para envío de spam,virus.

Aunque estas causas son las más comunes, existen muchas otras. Por lo que le sugerimos que, para su seguridad y la del resto de usuarios de Internet, revise su configuración, así como que instale, si aún no lo tiene, algún tipo de software antivirus, antiespías y/o cortafuegos en sus equipos, manteniéndolo actualizado con las últimas firmas de virus.

Si aún no dispone de este software, Movistar le sugiere:
- visitar la página de INTECO (Instituto Nacional de Tecnologías de la Comunicación), donde podrá encontrar herramientas de uso gratuito para el análisis de sus sistemas, a las que podrá acceder a través de la siguiente dirección:
http://cert.inteco.es/software/Proteccion/utiles_gratuitos/
- consultar el enlace para la gestión de incidencias de INTECO http://www.inteco.es/Seguridad
- el uso de paquetes de seguridad informática de cualquier proveedor.

Finalmente le recordamos que puede seguir enviando sus correos a través de:
- sus cuentas @telefonica.net, @terra.es, @infonegocio.com
- webmail de cualquier proveedor de correo
- sus cuentas de cualquier proveedor de correo si utiliza puerto SMTP seguro (ej: gmail, yahoo, ?)

En espera de que esta situación se resuelva lo antes posible.

Atentamente.

Madrid, 01 abr 2013

Nemesys Abuse Team
Telefonica de España S.A.U.

Figure 13 - Example ISP notification

11.2. CERT notification

This is an example of notification from INTECO-CERT to a hosting website contact. The website is involved on Stealrat botnet activity:

```
#################################################################
# INTECO-CERT is the security service of the National # Institute for Communication Technologies (INTECO) in Spain.
# Our main role is detection, coordination and response of # security incidents that take place on Spanish CI (Critical
# Infrastructure), Research and Academic Network (RedIRIS), # enterprises and/or citizens.
# Also we act as Spanish national CERT in the role of # coordination with other security teams.
#################################################################

Dear Team,

We have detected that the following website hosted on one of your servers is part of the Stealrat Botnet:

http://www.YYYY.com/wp-content/plugins/wysija-newsletters/models/menu.php

This website is hosted on your server with IP 46.16.62.xxx

You can find more information about how to detect the malicious files of this botnet in the following link:

http://blog.trendmicro.com/trendlabs-security-intelligence/how-to-check-if-your-website-is-part-of-the-stealrat-botnet/

In addition, stealrat botnet is analyzed in-depth in the following paper:

http://www.trendmicro.co.uk/media/wp/stealrat-whitepaper-en.pdf

Please check it and request the owner of the affected site to fix the issue and take appropiate update and security
measures so it doesn't happen again.

Please keep the incident tracking code in the subject line of your email to help us to keep track the incident.

Thank you.
Best Regards,

--
INTECO-CERT <incidencias@cert.inteco.es> PGP keys: http://www.inteco.es/what_is_inteco/About/PGP_Public_keys/
National Institute for Communication Technologies (INTECO) Avenida José Aguado, 41. Edificio INTECO
24005 León (Spain)
```

Figure 14 - Example CERT notification

11.3. Advisory in NSC of a spam campaign

SPAM advisory in German NSC:
http://blog.botfrei.de/2014/01/adobe-users-danger-spam-campaigns/

Figure 15 - Example advisory in NSC of a spam campaign - botfrei

The following content describes a SPAM campaign affecting Spanish Bank:
https://www.osi.es/es/actualidad/avisos/2014/05/detectada-oleda-de-correos-fraudulentos-que-suplantan-el-banco-popular

Figure 16 - Example advisory in NSC of spam campaign - OSI

11.4. Information about a specific BOTNET in NSC

Advisory about GameOver Zeus botnet in Germany NSC, and how to check:
http://blog.botfrei.de/2014/06/infected-gameover-zeus-perform-online-check/

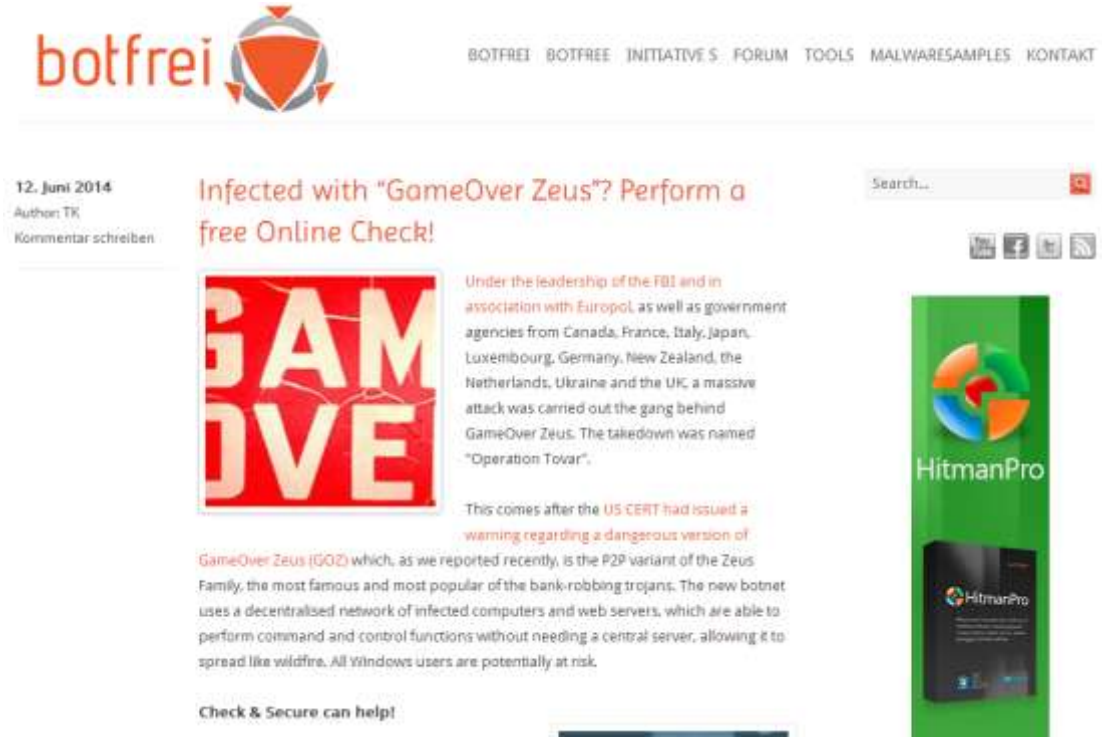

Figure 17 - Example information about a specific botnet in NSC - botfrei

The following content explains the spam botnet Pushdo, in the Spanish National Support Center:
http://www.osi.es/es/servicio-antibotnet/info/pushdo

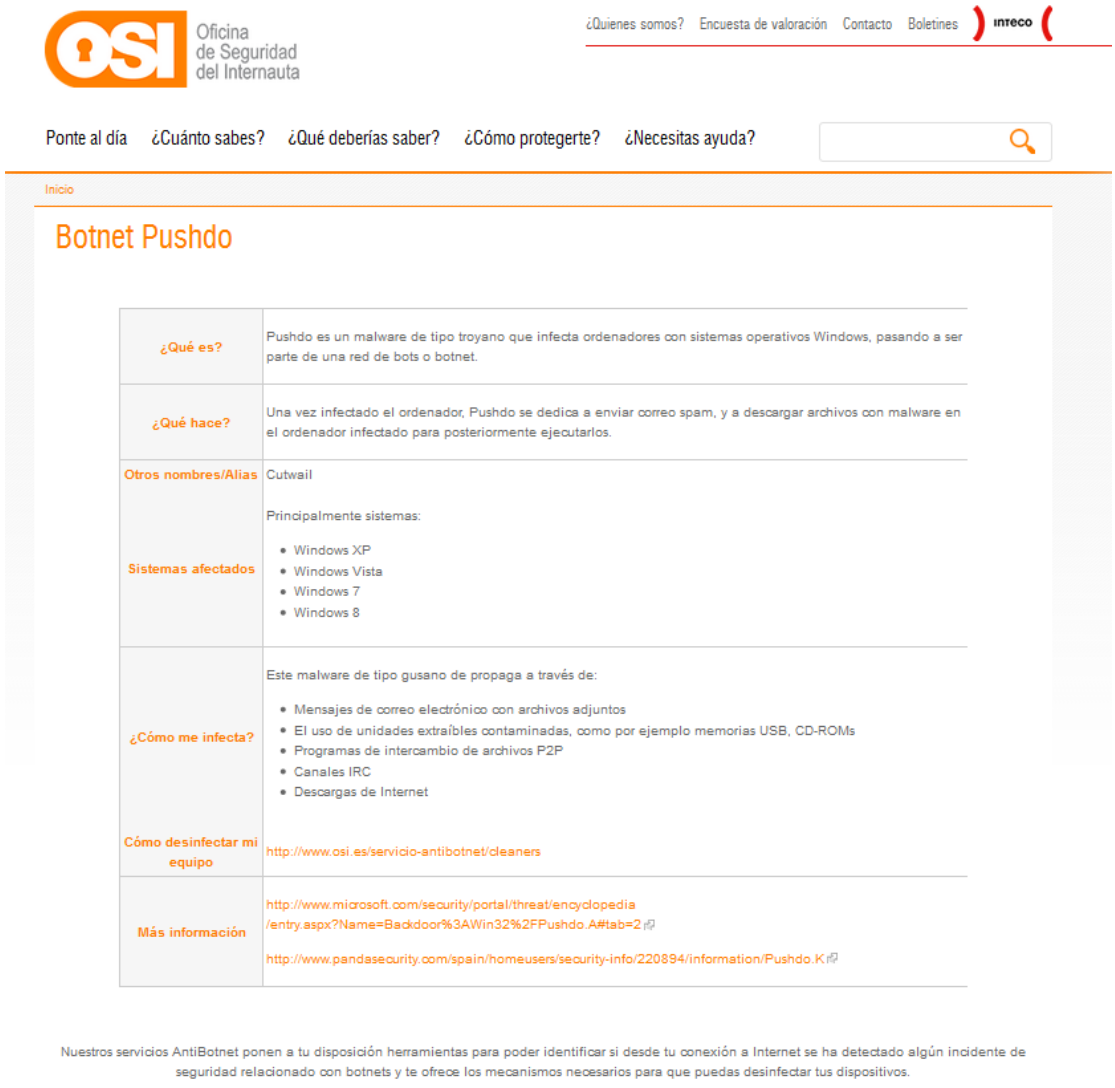

Figure 18 - Example information about a specific botnet in NSC - OSI

11.5. Advisory of a malicious APK in NSC

The following content inform Spanish end-users about a malicious APK in Android market:
https://www.osi.es/es/actualidad/avisos/2014/06/kk-tuneup-master-una-aplicacion-potencialmente-peligrosa

Figure 19 - Example advisory of malicious APK in NSC

11.6. Online "bot" checking service

The following service in Spanish NSC checks if from the user internet connection botnet activities have been identified:
https://www.osi.es/es/servicio-antibotnet

Figure 20 - Example online bot checking service - OSI

Figure 21 - Example online bot checking service - OSI

The Plugin for Chrome performs the checking periodically and alert the end-user:

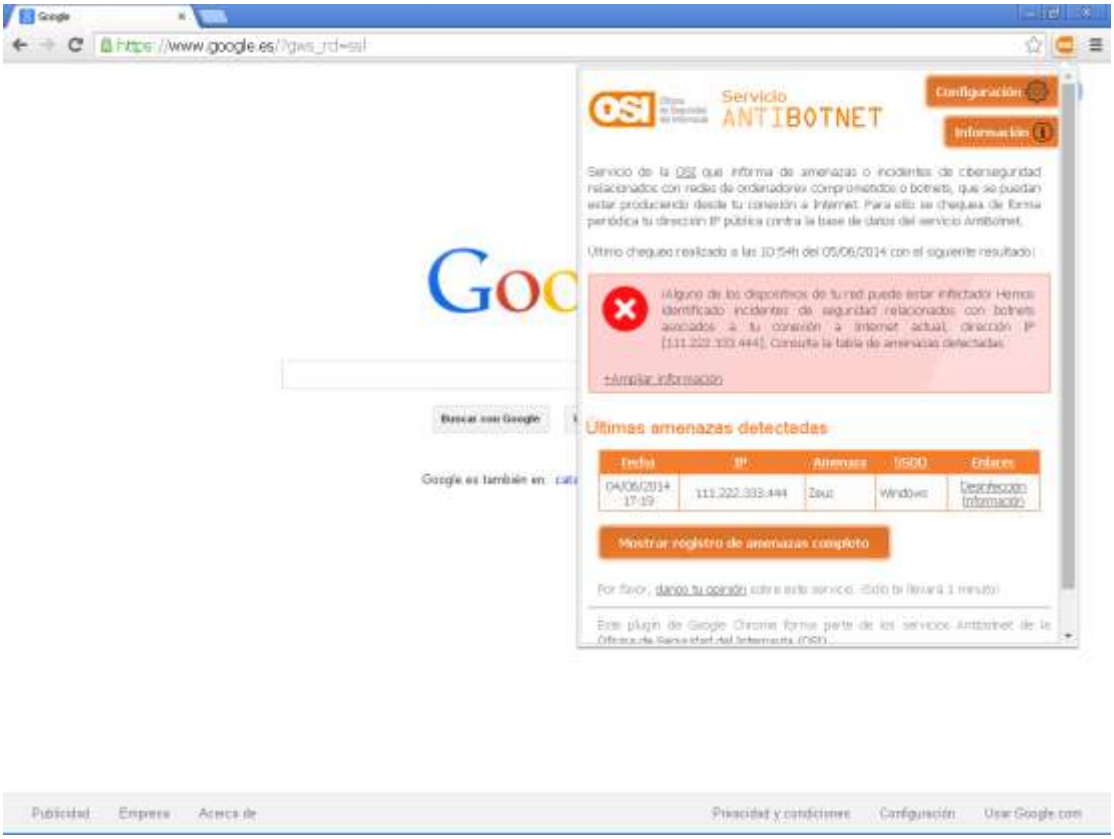

Figure 22 - Example online bot checking service - OSI

11.7. Cleaners recommendation

The following content provides free tools for bot disinfection:

Germany NSC:
https://www.botfrei.de/en/eucleaner.html

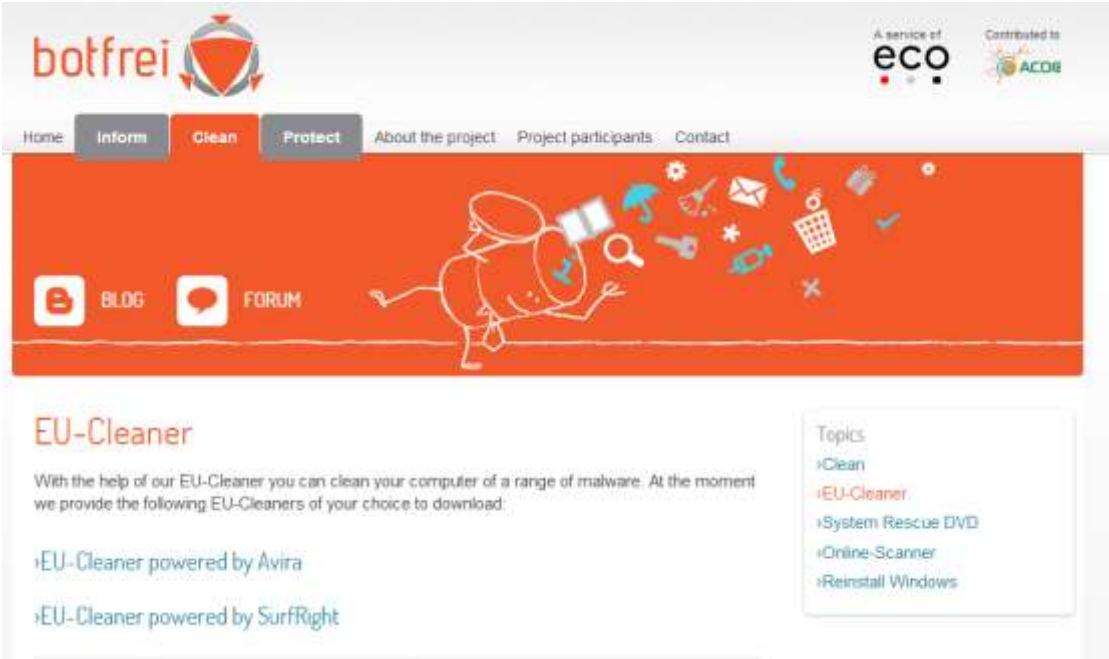

Figure 23 - Example cleaners recommendation - botfrei

Spanish NSC:
http://www.osi.es/es/servicio-antibotnet/cleaners

Figure 24 - Example cleaners recommendation - OSI