HORIZON 2020

Digital Security: Cybersecurity, Privacy and Trust
H2020-DS-2015-1

DS-04-2015 Information driven Cyber Security Management
Grant n° 700176

Secure Information Sharing Sensor Delivery event Network[†]

Deliverable D2.6: Metrics specification and results

Abstract: This deliverable presents the final set of metrics derived from the data available to the consortium and their application for cross-country and cross-region comparisons. With interactive and searchable online metrics & statistical mapping.

Contractual Date of Delivery	31 January 2019
Actual Date of Delivery	7 May 2019
Deliverable Security Class	Public
Editor	Will Rogofsky (CYBE)
Contributors	All *SISSDEN* partners
Quality Assurance	CYBE, NASK

[†] The research leading to these results has received funding from the European Union Horizon 2020 Programme (H2020-DS-2015-1) under grant agreement n° 700176.

The *SISSDEN* consortium consists of:

Naukowa i Akademicka Sieć Komputerowa	Coordinator	Poland
Montimage EURL	Principal Contractor	France
CyberDefcon Limited	Principal Contractor	United Kingdom
Universitaet des Saarlandes	Principal Contractor	Germany
Deutsche Telekom AG	Principal Contractor	Germany
Eclexys SAGL	Principal Contractor	Switzerland
Poste Italiane – Società per Azioni	Principal Contractor	Italy
Stichting the Shadowserver Foundation Europe	Principal Contractor	Netherlands

Table of Contents

1 Introduction

1.1 Aim of the document

The document presents the results of Task 2.5 “Metric development”. The aim of this task is to develop metrics that can be used to assess the effectiveness of the project and to quantify the level of attacks detected.

1.2 Work done so far

Since SISSDEN is an innovation project rather than a research project, the development of metrics based on SISSDEN data would depend on existing research. At the beginning of this task, the outcomes of the CyberROAD project[1] were studied to leverage its conclusions. In particular, CyberROAD performed a survey of existing cyber security indicators, but no suitable indicators could be found that assessed remediation rates. CyberROAD also produced a roadmap of research including areas of interest with respect to cyber security indicators.

Current research was then monitored, including in particular two deliverables (D3.1 and D3.2) in the SAINT project[2] which focused on an interesting area – a statistical model of calculating the cost benefits of information sharing between organisations. This work was assessed as to its suitability with respect to remediation rates – i.e. calculating the cost benefits of ingesting remediation reports from SISSDEN. However, the final deliverable was not due until December 2018 and ultimately could not be used, since the necessary data was not available.

Since no suitable research for accurately assessing the remediation benefits (and fulfilling the technical KPI for remediation rate) could be found, an alternative approach was decided to split the work into a qualitative assessment (via a survey distributed to report recipients) and quantitative assessment (via metrics of observed attacks). The survey was produced and results analysed, and the quantitative metrics published on an online dashboard.

1.3 Structure of the document

- Section 2 presents the objectives of the metrics and the challenges involved.
- Section 3 lists the requirements of the metrics, the survey and the online dashboard.
- Section 4 specifies the method of calculation of the metrics.
- Section 5 presents a (limited) set of results of the metrics, with the full results available online, and the full results of the survey.
- Section 6 presents the functionality of the online dashboard.
- Section 7 provides conclusions.

[1] https://www.cyberroad-project.eu/

[2] https://project-saint.eu/

2 Objectives

According to the Grant Agreement, the objectives of metrics in SISSDEN are twofold:

> *"to establish the scale of most important security issues in the EU, and impact of the project itself"*

The former objective (herein *Objective A*) aims to provide important data on the amount of security incidents recorded by SISSDEN within the EU. The latter objective (herein *Objective B*) aims to provide measurements of the impact that SISSDEN's outputs and results have had within the EU.

Here we discuss the motivation of these two objectives in further detail in the context of existing metrics and published research.

2.1 Objective A

Establish the scale of most important security issues in the EU

The aim of this objective is to provide data on the amount of security incidents recorded by SISSDEN, with a particular focus on those incidents from and within the EU. This information is useful primarily for the following purposes:

2.1.1 A.1: Verify existing research

Guidance in the form of regular (usually quarterly or annual) reports on the latest and most common threats are widely used by organisations, with some of the most popular being:

- ENISA Threat Landscape Report[3]
- McAfee Labs Threats Report[4]
- Imperva Web Application Attack Report[5]
- HPE Cyber Risk Report[6]
- Symantec Internet Security Threat Report[7]
- Cisco Annual Cybersecurity Report[8]
- FireEye Annual Threat Report[9]

In addition, there are several academic publications into the scale of cybercrime, but these are much more limited in availability and breadth of data than threat reports.

SISSDEN metrics can help verify existing research and determine which security issues may be currently overestimated or underestimated.

[3] https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018

[4] https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-dec-2018.pdf

[5] https://www.imperva.com/blog/the-state-of-web-application-vulnerabilities-in-2018/

[6] https://www.hpe.com/us/en/newsroom/blog-post/2017/03/hpe-cyber-risk-report-explains-cybersecurity-challenges-for-businesses.html

[7] https://www.symantec.com/security-center/threat-report

[8] https://www.cisco.com/c/m/en_au/products/security/offers/annual-cybersecurity-report-2018.html

[9] https://www.fireeye.com/current-threats/annual-threat-report.html

2.1.2 A.2: Prioritise anti-cybercrime expenditure

The reports listed in *Objective A.1* provide useful information to organisations at an executive level. In particular, they enable informed purchasing decisions for cybersecurity product licenses and service contracts.

However, the neutrality of these reports is often questionable, since they are written by commercial providers of cybersecurity solutions. A notable exception is the ENISA Threat Landscape Report, although even in this case the data is sourced from non-neutral parties (e.g. the 2018 report uses data provided by CYjAX[10]).

SISSDEN metrics can improve on this by offering a more neutral perspective. The core SISSDEN platform does not provide any products that prevent or mitigate cybercrime, but only detect and report. It is possible that SISSDEN results may be successfully exploited commercially (as described in D2.9), but in any case, the SISSDEN non-profit will continue to exist alongside any commercial solution.

2.1.3 A.3: Prioritise training and awareness

The reports listed in *Objective A.1* help stakeholders in the industry to remain up to date on the scale of the latest threats, and any newly emerging threats. They are particularly useful to SOC operatives, QA testers and software verification & validation engineers.

As with *Objective A.2*, SISSDEN metrics can improve on this by offering a more neutral perspective. In addition, SISSDEN metrics can be published on an online dashboard that enables the metrics to be interactively explored, which will enable more bespoke information for the relevant users.

2.2 Objective B

Establish the impact of the SISSDEN project in the EU

The aim of this objective is to provide a measure of impact that the SISSDEN project has achieved. This can either be in terms of the absolute numbers of remediations as a result of SISSDEN's reports, the percentage number of remediations, or the associated cost saving of the remediations. This would primarily be of interest to the European Commission as a measure of the direct effectiveness of the SISSDEN project.

In order to assess the impact, statistics of at least one of two types of data is required:

1. *Remediation statistics:* This data is required. An example of this data is the number of infected URLs and IPs resolved within X number of days. If used in isolation, this could be used to determine the number of incidents that SISSDEN is responsible for helping to remediate.
2. *Cost statistics:* This data is optional. By applying cost statistics on top of remediation statistics, estimates could be made as to how much money SISSDEN as a project has saved - either individual remediation report users, or across the EU.

Remediation statistics

Since remediation statistics are required, the initial question is how can SISSDEN obtain this information? Two methods can be considered:

[10] https://www.cyjax.com/

1. *External observation:* All incidents recorded by SISSDEN which make their way through to the free remediation reports could be compared to later reports to determine which URLs/IPs have been remediated.
2. *Third-party observation:* Individual users of the free remediation reports could provide data to SISSDEN that enables remediation rates to be calculated for the subset of reports that are sent to them.

The suggested approach for external observation is limited to the impact of the free remediation reports and discards any possible impact of other elements of SISSDEN (such as the analytical platform, published papers, blogs, etc). Despite limiting the scope in this way, it still has major weaknesses. Large assumptions would be required as to what determines a "remediated" incident. For example:

- How long after the incident is reported should be allowed for it to be removed from subsequent reports and classed as "remediated"?
- What happens if the same infection on the same URL/IP reappears after being classed as remediated? Should it be assumed that it wasn't remediated at all, and was possibly just offline or temporarily unreachable? Or should it be assumed that this is a new incident that happens to be of the same type on the same URL/IP?
- Even if the incident is truly remediated, it may not be remediated due to SISSDEN's reporting. It could be remediated via another source, or simply taken offline or moved elsewhere by the attackers.

It is clear that there are unacceptably large assumptions required to externally observe remediation rates without any internal confirmation. Therefore, the next option to be considered is a third-party approach. This would require users of the free remediation reports to send data back to SISSDEN on the number of reported incidents which have been remediated. In addition, they would have to report statistics for other remediations which have occurred during the same time period based on non-SISSDEN information; without this information, it would not be possible to determine whether a remediated incident was remediated solely due to a SISSDEN report or perhaps from multiple sources including SISSDEN.

This approach would be limited in scope to reports sent to particular companies. In other words, if 10 report recipients were to participate in this scheme, it would only be possible to declare with a high degree of confidence that, for example, Company A remediated 50% of SISSDEN reports, Company B remediated 60% of SISSDEN reports, etc. Extrapolating to the EU as a whole would result in low degrees of confidence. The larger the number of companies participating, the higher the degree of confidence. For example, when extrapolating the reported remediation rate to the EU as a whole, at least 40 average-sized report recipients would be required to partake in the scheme in order to achieve a reasonable margin of error (< 30% @ 95% confidence). Therefore, even if this limited scope can be accepted, there would be serious challenges in finding enough companies that can provide this data.

The SAINT project[11] (Deliverable 3.4, Section 2.3) reported the lack of metrics recorded by SOCs - only 54% of SOCs keep logs of any kind of metrics at all. More precisely and relevantly

[11] https://project-saint.eu/

in the case of SISSDEN, only 23% of SOCs keep logs of the "number of incidents handled" and only 20% record the "time from detection to containment or eradication". See *Figure 1*.

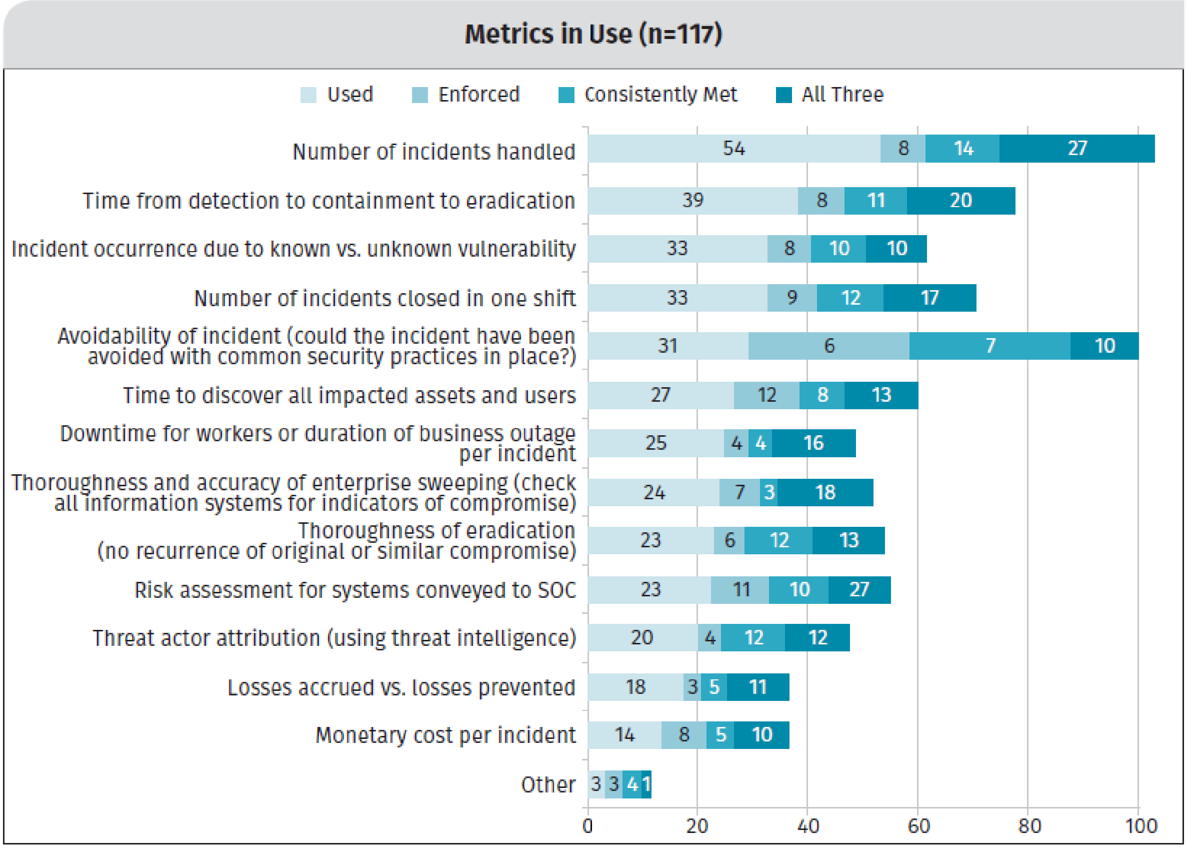


Figure 1: Metrics in use by SOCs

Cost statistics

The difficulties in obtaining accurate remediation data have been described above. Without remediation data, cost statistics (e.g. the average cost of damages caused by a DDoS attack) cannot be applied to number of remediations to produce an estimated cost saving.

In addition, there are difficulties in obtaining accurate cost data. The CyberROAD project[12] studied in detail the existing research in this field. Relevant findings from Deliverable 3.1[13], Section 3.2 include the following:

- Despite a large amount of data being available on cybercrime, the first systematic study of the costs of cybercrime did not take place until 2012 (*Measuring the Cost of Cybercrime*[14]).
- This study concludes that the available statistics on the costs of cybercrime are "insufficient and fragmented" and that the "lack of cohesion between different sources clouds the issue, leads to inconsistency of data and engenders mistrust of the numbers". Further, it states "this report supports the widely held opinion that despite eye-catching headlines suggesting otherwise, it remains the case that few

[12] https://www.cyberroad-project.eu/

[13] https://www.cyberroad-project.eu/m/filer_public/2016/05/02/d31_social_economic_political_and_legal_landscape_report.pdf

[14] https://link.springer.com/chapter/10.1007/978-3-642-39498-0_12

straightforward numbers exist on cybercrime and its true cost politically, economically, socially and morally".

- It also concludes that there is a lack of neutrality in most published cost statistics: "Previous studies of cybercrime have tended to study quite different things and were often written by organisations with an obvious 'agenda'".
- More recently than the 2012 study, CyberROAD also considers results from the Ponemon Institute's "Cost of Cyber Crime Study", which have been published annually since 2009. The Ponemon Institute is an independent research group, which helps overcome the neutrality issues of many other data providers. However, CyberROAD found that data either represents total organisational costs for all types of cybercrime, or that costs broken down by type of attack were limited by being applied to "difficult to assess categories" of attack, which were either not well-defined, too subjective or not easy to correlate to other events.

Several methodologies have been proposed which would enable more accurate calculation of cybercrime statistics. For example:

- *Measuring the Cost of Cybercrime* (published 2012) presents a model which would enable a calculation of the costs to society as a whole, which essentially resolves to the "sum of direct losses, indirect losses, and defence costs". However, this would require costs to be calculated on a per-company basis and SISSDEN could not find any current applications of this methodology.
- *Beyond the pretty penny: the Economic Impact of Cybercrime*[15] (published 2017) builds upon previous work including the aforementioned 2012 study.

Survey

An alternative approach to quantitatively measuring the remediation rate in an automated method is to qualitatively assess it instead. Although this would not solve the issues presented in this section of how to assert with certainty that an incident was remediated by SISSDEN, the advantage of a qualitative assessment is that other useful streams of information can also be gathered.

For this reason, a survey delivered to SISSDEN remediation report recipients would provide a useful method of asking users how many SISSDEN reports are successfully remediated by their company, but also discovering other aspects of SISSDEN's services that can be improved in the future.

2.3 Summary

In *Section 2.2*, it was demonstrated that the direct impact of remediations from SISSDEN is very complicated and in some sense impossible to accurately quantify:

- Remediation statistics cannot realistically be used directly from consumers of the free remediation reports, since this data is not widely recorded or available.
- Remediation statistics cannot be observed externally by SISSDEN, as several large assumptions would be required, to the point that the accuracy of the statistics would be extremely low.

[15] https://www.researchgate.net/publication/323372650

- Cost statistics do not use any well-defined categorisation of incidents and generally lack objectivity and neutrality, resulting in vastly inflated figures. More recently, methodologies have been proposed which attempt to solve these issues, but statistics based on these methodologies are not yet available.

Therefore, due to the issues with creating an automated, quantitative measure of the impact of the project, a reasonable alternative is to provide an ad-hoc, qualitative measure of the impact. For this purpose, SISSDEN will produce a survey which will find out how impactful SISSDEN's results have been for the relevant users.

Below, the Objectives that this section began with (taken from the GA) are repeated here with a revised summary of their intent. In addition, a third objective is created that covers the publications of the metrics.

2.3.1 Objective A

Report on the scale of the most important security issues in the EU

As detailed in *Section 2.1*, numbers of incidents recorded by SISSDEN over time will enable users to:

- A.1: Verify existing research
- A.2: Prioritise anti-cybercrime expenditure
- A.3: Prioritise training and awareness

The priority is to provide a set of metrics which enable SOCs to gain additional understanding into the most important and prevalent security issues in the EU from a source that is more independent and neutral than existing reporting providers. CyberROAD (in Deliverable 3.1, Section 3.2) demonstrated the bias and lack of neutrality from existing reporting providers.

2.3.2 Objective B

Establish the impact of the SISSDEN project in the EU

A survey will enable SISSDEN to receive qualitative information on the impact that SISSDEN has had on their remediation efforts. It will also enable vital feedback to be received on improving SISSDEN's future efforts and providing a more effective exploitation of the project.

2.3.3 Objective C

Publish the results on an online dashboard that enables the metrics to be effectively disseminated

As detailed in *Section 2.2*, results from the SAINT project were clear that the costs of cybercrime are extremely challenging for businesses to accurately quantify. As a result, SAINT highlighted that businesses should instead focus more strongly on cooperation and information sharing than quantification. For example, in Deliverable 3.1, Section 2.4, it was stated that *"Information sharing and interaction across organisations at the business level, provides to a certain degree, an economic incentive to support and promote the development of co-operation and to give priority to building relevant co-operational schemes."*

For this reason, SISSDEN should provide an online dashboard that enables a wide range of users to view and receive the relevant metrics. It will also enable the survey results to be distributed to a wide audience, and provide transparency into the answers to these questions.

A strong focus here should be building the dashboard with a modern user interface so that it is simple to use and able to engage a large range of users.

3 Requirements

The main requirements for the metrics are listed in the following sections sorted by subject and indicating their priority.

The requirement priority levels used are based on MoSCoW prioritization:

- MUST
- SHOULD
- COULD
- WON'T

Overall project requirements are listed in D3.1 Use Cases and Requirements. Requirements in this section, however, are specific to the metrics and have been selected in order to fulfill the objectives listed in *Section 2*.

3.1 Metrics

ID	Requirement	Priority	Associated objectives
1	The project will produce "absolute metrics" based on the absolute numbers of incidents detected by the platform.	MUST	A
2	The project will produce "proportional metrics" based on the normalized levels of incidents detected by the platform.	SHOULD	A
3	The project will use IP addresses as a baseline for normalization.	SHOULD	A
4	The project will produce metrics refined by threat category.	MUST	A
5	The project will produce metrics refined by geographic location.	SHOULD	A

3.2 Survey

ID	Requirement	Priority	Associated objectives
6	The survey will be aimed primarily at recipients of the remediation reports.	MUST	B
7	The survey will include a question on the approximate	MUST	B

	remediation rate achieved by users of SISSDEN's data.		
8	The survey will include additional questions to gauge the level of usefulness that respondents find SISSDEN's different types of data to be.	SHOULD	B
9	The survey will include an option to retain respondents' contact information in order to gather further feedback in the future.	SHOULD	B

3.3 Dashboard

ID	Requirement	Priority	Associated objectives
10	The metrics dashboard will be centred around a map view that draws immediate visual attention to a geographic representation of the metrics.	SHOULD	A, C
11	The metrics dashboard will enable metrics to be easily filtered by threat category.	SHOULD	A, C
12	The metrics dashboard will display an overview page of the metrics for each country.	SHOULD	A, C
13	The metrics dashboard will display pertinent results from the survey, with information redacted and anonymized appropriately.	SHOULD	B, C
14	The metrics dashboard will include suitable links to the SISSDEN website and the customer portal.	MUST	C
15	The metrics dashboard will provide low-level data that could potentially leak sensitive information from the sensor network to the public.	WON'T	A, C

4 Specification

In order to fulfill Objective A, two types of metrics will be produced in order to meet Requirement 1 & 2: absolute and proportional metrics. These metrics will represent the number of incidents detected by the platform and will take the following forms:

- Absolute metric

  4,230 DDoS amplification attacks detected from Sweden

- Proportional metric

  Malicious score of 45.4 out of 100 for DDoS amplification attacks detected from Sweden

In this context, "incidents detected by the platform" refers to all data that is collected by the platform and makes its way to the SISSDEN backend in the Elasticsearch cluster.

4.1 Categories

All data that is collected by the SISSDEN platform and makes its way to the SISSDEN backend in the Elasticsearch is eligible to be included in the metrics. Data that is located elsewhere (in partner or third-party deployments) is not available for ingestion.

It is not necessary to omit any data from the backend for reasons of privacy, since the metrics will aggregate to a high level (Requirement 14).

Elasticsearch Index	Description	Include	Category
amppot	DDoS amplification attacks from Amppot	No (lacks attacker IP)	-
badip	Attacking IPs	No (lacks attack categorisation)	-
cybe_daily_top_asn	Top ASNs for suspicious requests to CYBE darknet	No (aggregated to ASN)	-
cybe_daily_top_ip	Top IPs for suspicious requests to CYBE darknet	Yes	Suspicious Activity
cybe_daily_top_net	Top subnets for suspicious requests to CYBE darknet	No (aggregated to subnet)	-
nask_darknet*	Event data from NASK darknet	Yes	Suspicious Activity
nask_hpfeeds	External events contributed by T-Pot users	No (lacks attack categorisation)	-

nask_n6	Third-party data consolidated by NASK's n6	No (external data sources)	-
nask_pga*	PGA analysis results	No (contains only PGA analysis results)	-
nask_smtp	SMTP analysis results	No (sinkholed spam traffic)	-
usaar_iot	Attacks from USAAR IoT Lab	Yes	IoT Attacks
usaar_sandbox	Malware from USAAR sandbox analyses	No (lacks attacker IP)	-
ciscoasa-*	Attacks from Cisco ASA honeypot	Yes	Remote Code Execution
conpot	ICS attacks from Conpot	Yes	ICS Attacks
cowrie-*	SSH/Telnet attacks from Cowrie	Yes	Brute Force Attacks
dionaea-*	Attacks from Dionaea honeypot	Yes	Web & DB Attacks
elasticpot	Elasticsearch attacks from Elasticpot honeypot	Yes	Web & DB Attacks
glastopf	Web attacks from Glastopf honeypot	Yes	Web & DB Attacks
heralding-auth-*	Credentials from Heralding honeypot	No (credentials not needed)	-
heralding-sessions-*	Sessions from Heralding honeypot	Yes	Brute Force Attacks
honeypy-*	Attacks from HoneyPy honeypot	Yes	Web & DB Attacks
micros-*	Attacks from MICROS honeypot	Yes	Web & DB Attacks
rdpy-*	Remote desktop attacks from rdpy honeypot	Yes	RDP Attacks
spam-*	Captured spam mails	No (source IP not in own field, too computationally expensive to calculate at runtime)	-

struts-*	Attacks from Struts honeypot	Yes	Remote Code Execution
weblogic-*	Attacks from WebLogic honeypot	Yes	Remote Code Execution
malware-*	Captured malware metadata	No (not attack events)	-
pcaps_files	Metadata of PCAP files	No (not attack events)	-
pcaps_metrics	Metrics from PCAPs	No (not attack events)	-
pcaps_sessions-*	Session metadata from PCAPs	No (not attack events)	-
vps-nodes	Metadata of VPS nodes	No (not attack events)	-

Legend

Source of data

Partner systems	Sensor data	Non-attack/system data

This results in the following 7 categories of activity:

1. Brute Force Attacks
2. ICS Attacks
3. IoT Attacks
4. RDP Attacks
5. Remote Code Execution
6. Suspicious Activity
7. Web & DB Attacks

4.2 Normalisation

Since SISSDEN detects a very large number of attacks from a wide variety of attackers, it is necessary to normalise the results (Requirement 2). Since the metrics are also being categorised as a requirement, this will make it easier to compare across categories.

For example, imagine that large network is responsible for 5,000 brute force attacks that are detected by the SISSDEN platform, and likewise the network of a small company is responsible for 100. Which is “worse”? Certainly, it is too simplistic to say that the network responsible for 5,000 attacks is worse, simply because the absolute number of attacks is greater. The larger the network, the more attacks that would, on average, be expected to originate from the network.

To compensate for this, it is necessary to normalise for the “size” of the network. The obvious way to define a network is by ASN, but size is not so obvious. In reality, how “bad” a network is expected to be depends on a large range of parameters such as purpose of the network, infrastructure (bandwidth, number of devices, etc), number of users, type of user (constituents, employees, etc), company revenue, etc. The list of factors is potentially very long and clearly not information that is publicly available for an automated calculation.

A typical measure used for the size of the network is the number of IPv4 addresses. Although this is not a perfect measure, it is publicly available, can be automated, and does provide a rough approximation of the size of a network. Firstly, this is because IPv4 addresses are not cheap and are low in supply, so companies do not tend to possess more IPv4 addresses than is necessary. Secondly, simply because the greater the functions that are carried on a network, the greater the IP space is generally required.

The most common approach to normalising by IP address is to simply divide the number of incidents detected by the number of IP addresses. For large, similarly-sized networks, this tends to work, but there can be some unwanted side-effects to using this approach. Firstly, if small networks with very low numbers of IP addresses have a very small spike in attacks, then the normalised metric will be very high. This would make the metric extremely sensitive to false positives. For example, a network with 256 IP address with 1 attack detected would have a normalised value of 1/256 ≈ 0.0039. This would be higher than a network with 65,536 IP addresses and 249 incidents detected (≈ 0.0038). If the 1 attack is a false positive, then this would be grossly unfair.

Instead, the approach will be to apply a Bayesian weighting to each proportional metric. This Bayesian weighting will be calculated across the whole of the population (for example, when calculating the proportional metric of an ASN, the Bayesian weighting will be applied by comparing the Bayesian weighting of that of every other ASN). This approach is based on prior research by CyberDefcon researchers in a series of reports on ranking the worst ASNs. A full methodology of this approach can be viewed in the Annexes of these reports. For example, see the World Hosts Report series published by HostExploit[16].

After normalising by IP address using this approach, the result is an index with a range of 0 to 1,000, with 0 being good and 1,000 being bad. Since it uses a Bayesian weighting, however, in practice an ASN or country will never have an index of 0; because a Bayesian weighting is a population-derived factor, it cannot be stated with certainty that something does not exist within a population when you have only observed a sample of the population.

This index can be created on any set of consistent values within a population. Therefore, it can be applied to individual categories of attack, multiple categories of attack and across different entities (populations) such as ASN or country.

4.3 Weightings

In order to compare ASNs or countries against each other, it is desirable to have a single metric that can be compared in scale, since this cannot be achieved by looking at absolute statistics of attacks. The normalised index (*Section 4.2*) will be used to create an index for each category of attack on a common scale (0 to 1,000). Because each of these indices uses a common scale, a simple weighted average can be used to calculate an overall index for the entity (ASN/country).

To achieve this, a set of weightings must be assigned to each category. This can be applied either per category (e.g. Brute Force Attacks or ICS Attacks) or per source (e.g. Cowrie Brute Force Attacks). Since the level of importance of data differs between honeypots of the same category, the logical choice is to choose a category per source.

[16] http://hostexploit.com/?p=whr-201309

For each source, a weighting between 1 and 5 is chosen, with 5 giving the proportional index 5 times the significance of 1. The weightings can be seen below:

Source	**Category**	**Source weighting**	**Category weighting**
Cowrie	Brute Force Attacks	3	3
Heralding		3	
Conpot	ICS Attacks	4	4
USAAR IoT	IoT Attacks	2	2
Rdpy	RDP Attacks	5	5
CiscoASA	Remote Code Execution	2	2
Struts		2	
WebLogic		2	
CYBE Darknet	Suspicious Activity	3	6
NASK Darknet		3	
Dionaea	Web & DB Attacks	3	14
Elasticpot		3	
Glastopf		4	
HoneyPy		2	
MICROS		2	

The category weighting is included in the table which is an effective weighting calculated by summing up the individual source weightings for that category. These weightings are necessarily subjective and have been chosen based on the significance of events recorded by the source, the number of events recorded by the source, and the general reliability of the source.

The total index is calculated as a weighted average as follows:

```
SISSDEN Index = SUM(weighting * index) / SUM(weighting)
```

4.4 Geographies

Origin vs destination

Each incident recorded by the platform has both an originating address and a destination address. For data collected by the sensor network, the destination address is that of the sensor itself, whereas the originating address is that of the malicious actor.

Although the sensors have been deliberately installed in a wide variety of locations, they still do not span every country, and in each country that they are located they cover only a small number of IP addresses compared to that of the country as a whole. Therefore, whilst the distribution of the attacks across the different locations of the sensors is of some interest, it only provides a very limited view.

For these reasons, of primary interest is the originating address, which can be used to determine the locations where most of the malicious activity is sourced from.

Registration vs estimation

When determining the geographical location of an IP address, there is no simple answer. Since there is no provable technical solution that can determine the location of an IP address, GeoIP services use a variety of deterministic factors to produce an estimation. However, the accuracy of these estimations still varies significantly. MaxMind (probably the biggest provider of commercial GeoIP services) tested the accuracy of their services by compared estimated locations with the real location answered by users[17]. The results show that the data is fairly accurate for large, technologically-developed countries, but less so for others. For example, IPs from the United States had an 86% accuracy level, whereas Venezuela had just 40%.

Due to this difficulty, generally when reporting data on IPs or ASNs, the "registered country" is used instead. This is the country of registration for the IP block or ASN as reported by the relevant Regional Internet Registry. In the results in the document, this is the country that is used for ASNs, likewise determines which ASNs are included within a country, and therefore influence the lists of top countries.

However, an additional and complementary approach will also be taken. An "estimated country" will be determined for each ASN. In most cases, this is likely to be the same as the registered country. But in the few cases where it differs from the registered country, then this information can be very interesting, even if it is not always for malicious reasons.

The method for determining the estimated country is based on prior research by CYBE, and is essentially determined by the following steps:

1. A list of all IP CIDR-formatted blocks is retrieved, as announced publicly over BGP. The Routeviews[18] project is used for this purpose, with all publicly announced routes being retrieved from a minimum of 3 locations.
2. Most specific route wins - even if a larger route is registered, it is disregarded if this is not reflected in public announcements.

[17] https://www.maxmind.com/en/geoip2-city-accuracy-comparison

[18] http://www.routeviews.org

3. GeoIP location of each route is retrieved from MaxMind.
4. Accuracy level from Maxmind of each route is aggregated to ASN level. Each ASN will have an accuracy level for multiple countries.
5. ASN-level accuracy is then combined with the number of routes that have conflicting announcements in their country and peering (between each route view). If any of these routes align with a GeoIP a country other than the registered country, this may increase its accuracy to a majority. At this point, this country is considered to be the estimated country.

4.5 Generation

- Metrics will are generated at 24-hour intervals and use the data from the previous 30 days in the backend. This period was selected after generating metrics over several periods. When including data for shorter periods of time (e.g. 24 hours), it was found that because some honeypots detect fewer but potentially more serious attacks, they did not have a large enough spread of attacks across IP space to draw conclusions. Longer periods (e.g. 90 days) meant that there was not a significant amount of variation in the results from one day to the next, even when a spike of attacks occurred, since each day when the metrics would be recalculated, 89 days of the data would remain stagnant. 30 days was selected as a “sweet spot”.
- Some indexes contain data derived from the IP address such as ASN, AS Organization etc. In some cases, using this data will be more accurate than calculating the ASN at the time of generation of the metrics, since the IP may have changed routing in this time. However, the effect is minimal, and not all indexes contain this data. Therefore, for consistency, the metrics fetch this info at the time of generation, using data from the closest time possible to the period of generation.

5 Results

- For brevity, the results listed below are limited to the top 10 results per query. The full results can be viewed on the Metrics Dashboard.
- Rows highlighted in blue represent countries in the EU.
- In this section, “Country” refers to “Registered Country”. In the context of an ASN, this represents the country of registration as reported by the relevant Regional Internet Registry. Similarly, lists of top countries are generated by including attacks from all ASNs registered to this country.

5.1 Absolute metrics

5.1.1 Brute Force Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	57043	HOSTKEY-AS	NL	13,568	41,254,739
2	14061	DIGITALOCEAN-ASN	US	1,863,680	38,625,252
3	57509	LL-INVESTMENT-LTD	BG	256	27,651,408
4	49453	GLOBALLAYER	NL	25,344	19,253,130
5	4134	CHINANET-BACKBONE	CN	115,909,632	19,020,439
6	16276	OVH	FR	3,102,976	10,819,020
7	60355	KVSOLUTIONSNL	NL	1,280	10,180,722
8	4837	CHINA169-BACKBONE	CN	58,292,480	9,708,868
9	54290	HOSTWINDS	US	305,408	9,666,772
10	56046	CMNET-JIANGSU-AP	CN	5,138,432	6,404,220

Top Countries

#	Country	Name	ASNs	IPs	Count
1	NL	Netherlands	896	32,114,304	84,279,871
2	US	United States	16,883	1,297,353,844	59,528,380
3	CN	China	422	440,711,680	42,609,490
4	BG	Bulgaria	632	6,040,832	27,798,657
5	FR	France	1,085	83,829,248	13,142,735
6	RU	Russian Federation	5,055	61,620,224	6,275,283
7	DE	Germany	1,857	120,647,296	4,437,352
8	BR	Brazil	6,246	154,998,400	3,356,441
9	UA	Ukraine	1,755	14,999,808	2,592,881
10	IN	India	1,692	46,648,920	2,214,277

5.1.2 ICS Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	45090	CNNIC-TENCENT-NET-AP	CN	4,953,856	211,029
2	10439	CARINET	US	121,600	45,421
3	8369	INTERSVYAZ-AS	RU	363,264	39,457
4	202425	INT-NETWORK	SC	277,760	35,390
5	63949	LINODE-AP	US	544,512	28,907
6	32475	SINGLEHOP-LLC	US	453,888	18,150
7	38365	CNNIC-BAIDU-AP	CN	620,032	15,424
8	209299	VITOX-TELECOM	NL	3,840	15,083
9	132203	TENCENT-NET-AP-CN	CN	781,824	14,647
10	55933	CLOUDIE-AS-AP	HK	203,520	14,504

Top Countries

#	Country	Name	ASNs	IPs	Count
1	CN	China	422	440,711,680	360,883
2	US	United States	16,883	1,297,353,844	159,785
3	RU	Russian Federation	5,055	61,620,224	49,599
4	HK	Hong Kong	497	25,690,924	47,189
5	SC	Seychelles	14	385,792	39,760
6	NL	Netherlands	896	32,114,304	32,608
7	TH	Thailand	362	19,793,024	23,341
8	KR	Korea, Republic of	737	159,708,288	20,438
9	ID	Indonesia	1,094	25,360,000	12,882
10	FR	France	1,085	83,829,248	8,116

5.1.3 IoT Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	202425	INT-NETWORK	SC	277,760	25,601
2	49981	WORLDSTREAM	NL	62,976	18,694
3	49505	SELECTEL	RU	296,960	16,156
4	204428	SS-NET	BG	256	9,217
5	57043	HOSTKEY-AS	NL	13,568	7,973
6	200391	KREZ999AS	BG	256	7,074
7	14061	DIGITALOCEAN-ASN	US	1,863,680	5,966
8	16276	OVH	FR	3,102,976	5,693
9	51852	PLI-AS	CH	30,720	3,939
10	4837	CHINA169-BACKBONE	CN	58,292,480	3,362

Top Countries

#	Country	Name	ASNs	IPs	Count
1	NL	Netherlands	896	32,114,304	30,119
2	RU	Russian Federation	5,055	61,620,224	30,104
3	SC	Seychelles	14	385,792	27,642
4	BG	Bulgaria	632	6,040,832	16,832
5	US	United States	16,883	1,297,353,844	16,769
6	CN	China	422	440,711,680	11,773
7	FR	France	1,085	83,829,248	6,080
8	CH	Switzerland	627	13,148,160	3,946
9	BR	Brazil	6,246	154,998,400	2,522
10	LV	Latvia	212	1,838,592	2,430

5.1.4 RDP Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	49505	SELECTEL	RU	296,960	6,072
2	16509	AMAZON-02	US	22,736,640	3,496
3	812	ROGERS-COMMUNICATIONS	CA	7,179,264	3,453
4	57271	BITWEB-AS	RU	512	2,733
5	16276	OVH	FR	3,102,976	2,534
6	50340	SELECTEL-MSK	RU	60,160	1,493
7	200391	KREZ999AS	BG	256	1,144
8	20473	AS-CHOOPA	US	963,840	1,004
9	43350	NFORCE	NL	78,848	846
10	60798	ASSERVEREASY	IT	4,096	672

Top Countries

#	Country	Name	ASNs	IPs	Count
1	RU	Russian Federation	5,055	61,620,224	11,041
2	US	United States	16,883	1,297,353,844	6,371
3	CA	Canada	1,250	74,713,344	3,474
4	FR	France	1,085	83,829,248	2,782
5	NL	Netherlands	896	32,114,304	1,903
6	BG	Bulgaria	632	6,040,832	1,175
7	CN	China	422	440,711,680	1,085
8	IT	Italy	884	55,023,872	724
9	VN	Viet Nam	244	40,048,128	541
10	TH	Thailand	362	19,793,024	407

5.1.5 Remote Code Execution

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	44050	PIN-AS	RU	213,760	2,056
2	63949	LINODE-AP	US	544,512	573
3	14061	DIGITALOCEAN-ASN	US	1,863,680	557
4	27699	TELEFÔNICA	BR	6,607,104	504
5	6939	HURRICANE	US	602,624	364
6	206791	SBY-TELECOM-AS	UA	256	355
7	45090	CNNIC-TENCENT-NET-AP	CN	4,953,856	353
8	10439	CARINET	US	121,600	342
9	4134	CHINANET-BACKBONE	CN	115,909,632	312
10	237	MERIT-AS-14	US	5,806,336	223

Top Countries

#	Country	Name	ASNs	IPs	Count
1	US	United States	16,883	1,297,353,844	3,916
2	RU	Russian Federation	5,055	61,620,224	3,085
3	CN	China	422	440,711,680	1,791
4	BR	Brazil	6,246	154,998,400	1,211
5	UA	Ukraine	1,755	14,999,808	717
6	IN	India	1,692	46,648,920	432
7	GB	United Kingdom	1,847	110,815,296	407
8	FR	France	1,085	83,829,248	332
9	TR	Turkey	441	29,948,672	318
10	IR	Iran, Islamic Republic of	441	20,930,304	300

5.1.6 Suspicious Activity

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	136782	PINGTAN-AS-AP	CN	11,008	7,044,864
2	209299	VITOX-TELECOM	NL	3,840	652,985
3	14061	DIGITALOCEAN-ASN	US	1,863,680	342,473
4	45899	VNPT-AS-VN	VN	15,923,456	284,366
5	201912	FCLOUD-AS	DE	512	214,144
6	204428	SS-NET	BG	256	209,903
7	202242	ARUBA-CLOUD	IT	26,624	191,555
8	204875	ISTEC-AS	UA	256	131,567
9	38265	SSKRU-AS-AP	TH	512	116,797
10	49981	WORLDSTREAM	NL	62,976	113,483

Top Countries

#	Country	Name	ASNs	IPs	Count
1	CN	China	422	440,711,680	7,047,689
2	NL	Netherlands	896	32,114,304	951,708
3	US	United States	16,883	1,297,353,844	586,931
4	VN	Viet Nam	244	40,048,128	362,094
5	DE	Germany	1,857	120,647,296	222,571
6	BG	Bulgaria	632	6,040,832	216,121
7	IT	Italy	884	55,023,872	200,125
8	HK	Hong Kong	497	25,690,924	197,474
9	RU	Russian Federation	5,055	61,620,224	176,645
10	UA	Ukraine	1,755	14,999,808	145,603

5.1.7 Web & DB Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Count
1	56041	CMNET-ZHEJIANG-AP	CN	5,521,920	16,733,530
2	4837	CHINA169-BACKBONE	CN	58,292,480	14,757,572
3	56300	MYREPUBLIC-SG	SG	32,256	14,587,436
4	56046	CMNET-JIANGSU-AP	CN	5,138,432	12,704,778
5	9808	CMNET-GD	CN	54,299,392	11,555,923
6	15169	GOOGLE	US	12,921,088	9,788,204
7	4134	CHINANET-BACKBONE	CN	115,909,632	7,799,517
8	9304	HUTCHISON-AS-AP	HK	4,102,400	6,526,094
9	51659	ASBAXET	RU	8,960	5,148,094
10	136190	CHINATELECOM-YUNNAN-DALI-MAN	CN	67,072	4,871,616

Top Countries

#	Country	Name	ASNs	IPs	Count
1	CN	China	422	440,711,680	81,213,633
2	US	United States	16,883	1,297,353,844	17,420,207
3	SG	Singapore	329	8,747,776	15,052,904
4	HK	Hong Kong	497	25,690,924	6,773,853
5	RU	Russian Federation	5,055	61,620,224	5,578,999
6	UA	Ukraine	1,755	14,999,808	4,915,810
7	CO	Colombia	135	18,810,496	4,811,664
8	ZA	South Africa	385	41,974,784	3,475,461
9	JM	Jamaica	8	358,656	2,703,708
10	FR	France	1,085	83,829,248	1,442,013

5.2 Proportional metrics

5.2.1 Brute Force Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	14061	DIGITALOCEAN-ASN	US	1,863,680	969.4
2	57043	HOSTKEY-AS	NL	13,568	950.4
3	60355	KVSOLUTIONSNL	NL	1,280	920.7
4	54290	HOSTWINDS	US	305,408	919.2
5	49981	WORLDSTREAM	NL	62,976	912.6
6	53667	PONYNET	US	58,368	903.8
7	46664	VDI-NETWORK	US	9,728	903.7
8	197226	SPRINT-SDC	PL	15,872	901.9
9	60134	AS-STARTNIX	ES	2,816	901.6
10	51659	ASBAXET	RU	8,960	901.4

Top Countries

#	Country	Name	ASNs	IPs	Index
1	NL	Netherlands	896	32,114,304	1,000.0
2	BG	Bulgaria	632	6,040,832	486.0
3	US	United States	16,883	1,297,353,844	191.1
4	CN	China	422	440,711,680	186.7
5	FR	France	1,085	83,829,248	177.2
6	AL	Albania	62	614,144	164.9
7	LV	Latvia	212	1,838,592	157.7
8	SC	Seychelles	14	385,792	151.1
9	RU	Russian Federation	5,055	61,620,224	147.8
10	UA	Ukraine	1,755	14,999,808	126.1

5.2.2 ICS Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	209299	VITOX-TELECOM	NL	3,840	907.1
2	10439	CARINET	US	121,600	527.1
3	200651	FLOKINET	SC	3,584	336.5
4	138415	HENGDA-HK	HK	16,128	305.4
5	202425	INT-NETWORK	SC	277,760	267.0
6	45090	CNNIC-TENCENT-NET-AP	CN	4,953,856	253.6
7	8369	INTERSVYAZ-AS	RU	363,264	248.9
8	55967	CNNIC-BAIDU-AP	CN	44,032	220.9
9	132701	URU-AS	TH	512	219.1
10	38044	GITN-NETWORK	MY	4,608	195.0

Top Countries

#	Country	Name	ASNs	IPs	Index
1	SC	Seychelles	14	385,792	911.1
2	CN	China	422	440,711,680	222.7
3	HK	Hong Kong	497	25,690,924	162.5
4	US	United States	16,883	1,297,353,844	148.1
5	TH	Thailand	362	19,793,024	138.0
6	NL	Netherlands	896	32,114,304	136.6
7	RU	Russian Federation	5,055	61,620,224	135.9
8	ID	Indonesia	1,094	25,360,000	117.4
9	KR	Korea, Republic of	737	159,708,288	109.3
10	MY	Malaysia	182	14,242,560	109.1

5.2.3 IoT Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	204428	SS-NET	BG	256	936.0
2	200391	KREZ999AS	BG	256	741.6
3	49981	WORLDSTREAM	NL	62,976	569.1
4	57043	HOSTKEY-AS	NL	13,568	548.7
5	35582	CHISTYAKOV	RU	256	371.2
6	202425	INT-NETWORK	SC	277,760	351.2
7	49877	RMINJINERING	RU	1,024	308.3
8	57271	BITWEB-AS	RU	512	296.2
9	41390	RN-DATA-	LV	1,792	261.6
10	200651	FLOKINET	SC	3,584	260.2

Top Countries

#	Country	Name	ASNs	IPs	Index
1	SC	Seychelles	14	385,792	994.1
2	BG	Bulgaria	632	6,040,832	253.3
3	NL	Netherlands	896	32,114,304	236.5
4	RU	Russian Federation	5,055	61,620,224	219.3
5	US	United States	16,883	1,297,353,844	157.6
6	LV	Latvia	212	1,838,592	143.2
7	CN	China	422	440,711,680	141.2
8	EE	Estonia	104	1,310,464	127.3
9	CH	Switzerland	627	13,148,160	124.7
10	FR	France	1,085	83,829,248	123.6

5.2.4 RDP Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	57271	BITWEB-AS	RU	512	945.0
2	200391	KREZ999AS	BG	256	458.0
3	49505	SELECTEL	RU	296,960	315.0
4	60798	ASSERVEREASY	IT	4,096	278.5
5	50340	SELECTEL-MSK	RU	60,160	236.4
6	57043	HOSTKEY-AS	NL	13,568	170.3
7	43350	NFORCE	NL	78,848	165.3
8	812	ROGERS-COMMUNICATIONS	CA	7,179,264	159.8
9	16509	AMAZON-02	US	22,736,640	158.5
10	16276	OVH	FR	3,102,976	146.6

Top Countries

#	Country	Name	ASNs	IPs	Index
1	RU	Russian Federation	5,055	61,620,224	1,000.0
2	BG	Bulgaria	632	6,040,832	872.7
3	EE	Estonia	104	1,310,464	480.8
4	SC	Seychelles	14	385,792	383.9
5	NL	Netherlands	896	32,114,304	379.1
6	CA	Canada	1,250	74,713,344	340.4
7	FR	France	1,085	83,829,248	274.6
8	TH	Thailand	362	19,793,024	194.3
9	US	United States	16,883	1,297,353,844	180.7
10	CO	Colombia	135	18,810,496	173.5

5.2.5 Remote Code Execution

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	51659	ASBAXET	RU	8,960	662.3
2	10439	CARINET	US	121,600	389.7
3	46805	CACHED	US	28,416	338.9
4	206791	SBY-TELECOM-AS	UA	256	305.9
5	63949	LINODE-AP	US	544,512	254.1
6	31214	TIS-DIALOG-AS	RU	58,368	223.7
7	209299	VITOX-TELECOM	NL	3,840	221.0
8	204428	SS-NET	BG	256	212.9
9	9542	TOM-AS-AP	MY	4,864	202.6
10	44050	PIN-AS	RU	213,760	200.5

Top Countries

#	Country	Name	ASNs	IPs	Index
1	SC	Seychelles	14	385,792	904.8
2	RU	Russian Federation	5,055	61,620,224	348.2
3	UA	Ukraine	1,755	14,999,808	284.0
4	US	United States	16,883	1,297,353,844	202.9
5	LR	Liberia	9	322,560	197.0
6	GT	Guatemala	37	2,095,104	186.5
7	BD	Bangladesh	649	2,464,256	176.7
8	CN	China	422	440,711,680	159.5
9	TR	Turkey	441	29,948,672	143.9
10	IS	Iceland	60	928,256	135.9

5.2.6 Suspicious Activity

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	14061	DIGITALOCEAN-ASN	US	1,863,680	969.4
2	57043	HOSTKEY-AS	NL	13,568	950.4
3	60355	KVSOLUTIONSNL	NL	1,280	920.7
4	54290	HOSTWINDS	US	305,408	919.2
5	49981	WORLDSTREAM	NL	62,976	912.6
6	53667	PONYNET	US	58,368	903.8
7	46664	VDI-NETWORK	US	9,728	903.7
8	197226	SPRINT-SDC	PL	15,872	901.9
9	60134	AS-STARTNIX	ES	2,816	901.6
10	51659	ASBAXET	RU	8,960	901.4

Top Countries

#	Country	Name	ASNs	IPs	Index
1	SC	Seychelles	14	385,792	900.9
2	BG	Bulgaria	632	6,040,832	632.9
3	NL	Netherlands	896	32,114,304	604.8
4	CN	China	422	440,711,680	472.0
5	UA	Ukraine	1,755	14,999,808	260.1
6	KH	Cambodia	81	951,552	256.8
7	VN	Viet Nam	244	40,048,128	256.7
8	HK	Hong Kong	497	25,690,924	230.7
9	TH	Thailand	362	19,793,024	208.5
10	ID	Indonesia	1,094	25,360,000	179.6

5.2.7 Web & DB Attacks

Top ASNs

#	ASN	AS name	Country	IPs	Index
1	209299	VITOX-TELECOM	NL	3,840	491.5
2	197226	SPRINT-SDC	PL	15,872	466.5
3	51659	ASBAXET	RU	8,960	319.5
4	15169	GOOGLE	US	12,921,088	303.7
5	60355	KVSOLUTIONSNL	NL	1,280	283.9
6	204915	AWEX	US	768	280.9
7	206804	ESTNOC-AS	EE	4,096	263.3
8	35067	PROKK-AS	UA	1,024	261.2
9	133774	CHINATELECOM-FUJIAN-FUZHOU-IDC1	CN	157,184	242.7
10	202242	ARUBA-CLOUD	IT	26,624	234.4

Top Countries

#	Country	Name	ASNs	IPs	Index
1	SC	Seychelles	14	385,792	398.3
2	CN	China	422	440,711,680	356.7
3	EE	Estonia	104	1,310,464	286.3
4	US	United States	16,883	1,297,353,844	233.6
5	FR	France	1,085	83,829,248	214.4
6	HK	Hong Kong	497	25,690,924	208.8
7	RU	Russian Federation	5,055	61,620,224	177.5
8	JM	Jamaica	8	358,656	176.1
9	ME	Montenegro	17	179,968	170.1
10	LR	Liberia	9	322,560	161.3

5.3 Survey

The survey was advertised on the website via a popup notice and distributed via mailing lists and social media. As of 10th April 2019, 149 valid responses were received. Aggregated results are shown below per question and a Summary section follows.

5.3.1 Question 1

In which country is your organisation primarily based?	
Type	Optional
Responses	144 (97%)

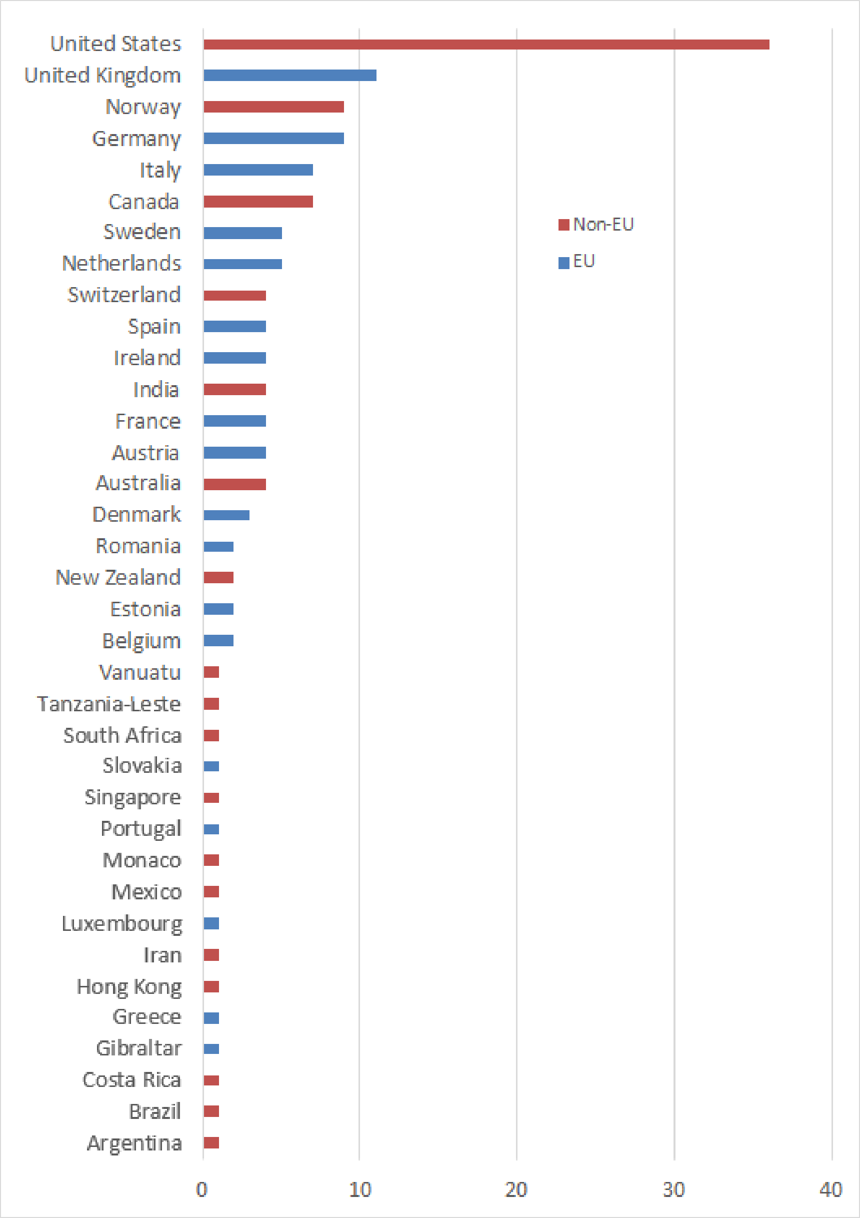


Figure 2: Respondent countries

5.3.2 Question 2

Approximately how many employees work for your organisation?	
Type	Optional
Responses	145 (97%)

Region	Response	Count	Percentage
EU	1-5	7	10.8%
	6-20	11	16.9%
	21-100	15	23.1%
	101-500	4	6.2%
	501-1,000	3	4.6%
	1,000+	25	38.5%
	Total	65	100%
Other	1-5	6	7.5%
	6-20	5	6.3%
	21-100	8	10.0%
	101-500	16	20.0%
	501-1,000	4	5.0%
	1,000+	41	51.3%
	Total	80	100%

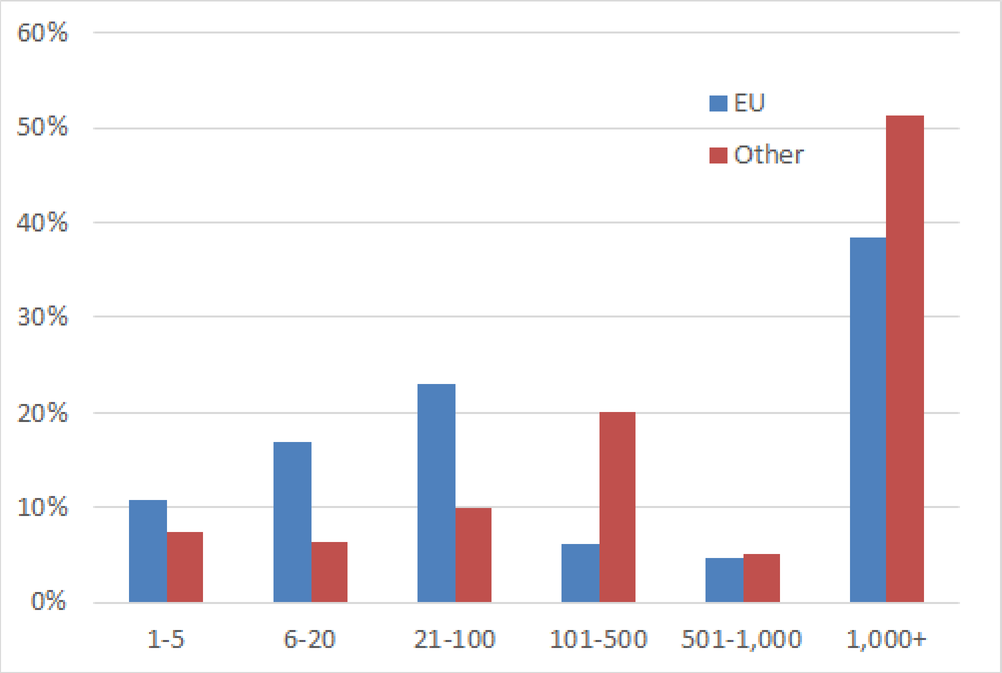


Figure 3: Respondents by company size

5.3.3 Question 3

Which category most closely fits your organisation type?	
Type	Optional
Responses	146 (98%)

Region	Response	Count	Percentage
EU	Academia/Research	17	25.8%
	Private enterprise	13	19.7%
	Internet service provider or operator	14	21.2%
	National CERT or CSIRT	8	12.1%
	Other	4	6.1%
	Web hosting provider or registrar	8	12.1%
	Policy making, government or legal	1	1.5%
	Independent cyber security practitioner or expert	1	1.5%

	Total	66	100%
Other	Academia/Research	20	25.0%
	Private enterprise	19	23.8%
	Internet service provider or operator	18	22.5%
	National CERT or CSIRT	8	10.0%
	Other	9	11.3%
	Web hosting provider or registrar	2	2.5%
	Policy making, government or legal	3	3.8%
	Independent cyber security practitioner or expert	1	1.3%
	Total	80	100%

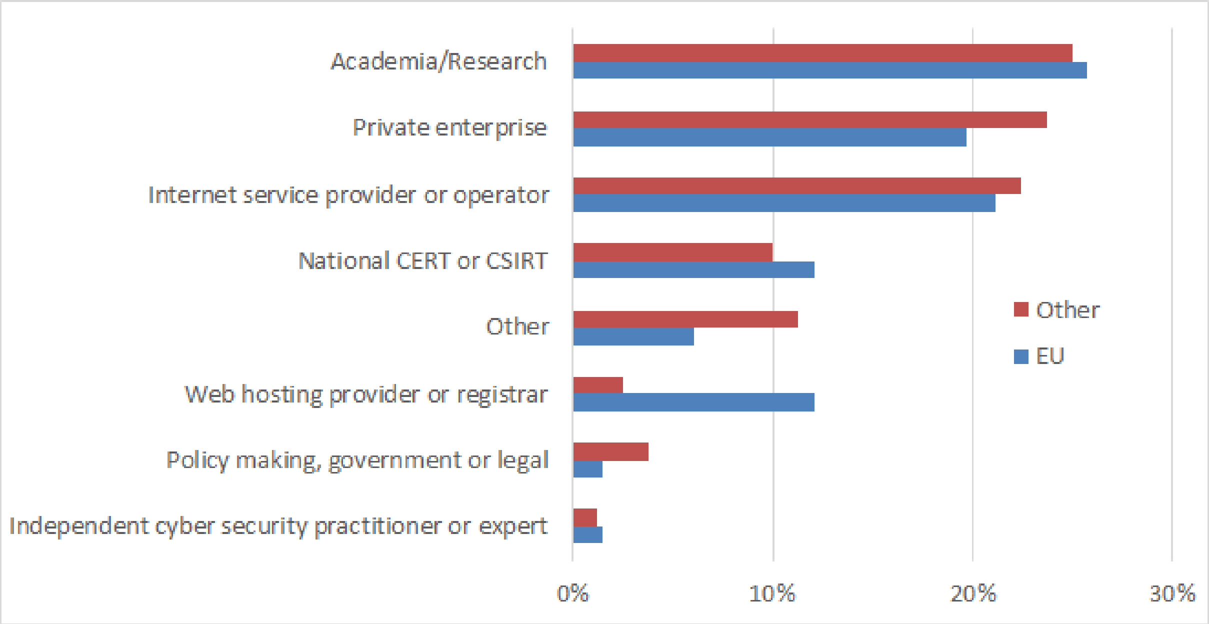


Figure 4: Respondents by organisation type

5.3.4 Question 4

Do you action remediation reports for your own end-users (e.g. in an enterprise) or for constituents (e.g. in a national CERT)? If your own constituents, please state how many constituents you serve.	
Type	Optional
Responses	143 (96%)

Region	Response	Count	Percentage
EU	Our own end users	54	85.7%
	10M+	1	1.6%
	1M-10M	2	3.2%
	100,000-1M	0	0.0%
	10,000-100,000	0	0.0%
	1,000-10,000	3	4.8%
	0-1,000	3	4.8%
	Total	63	100%
Other	Our own end users	75	93.8%
	10M+	0	0.0%
	1M-10M	0	0.0%
	100,000-1M	1	1.3%
	10,000-100,000	1	1.3%
	1,000-10,000	2	2.5%
	0-1,000	1	1.3%
	Total	80	100%

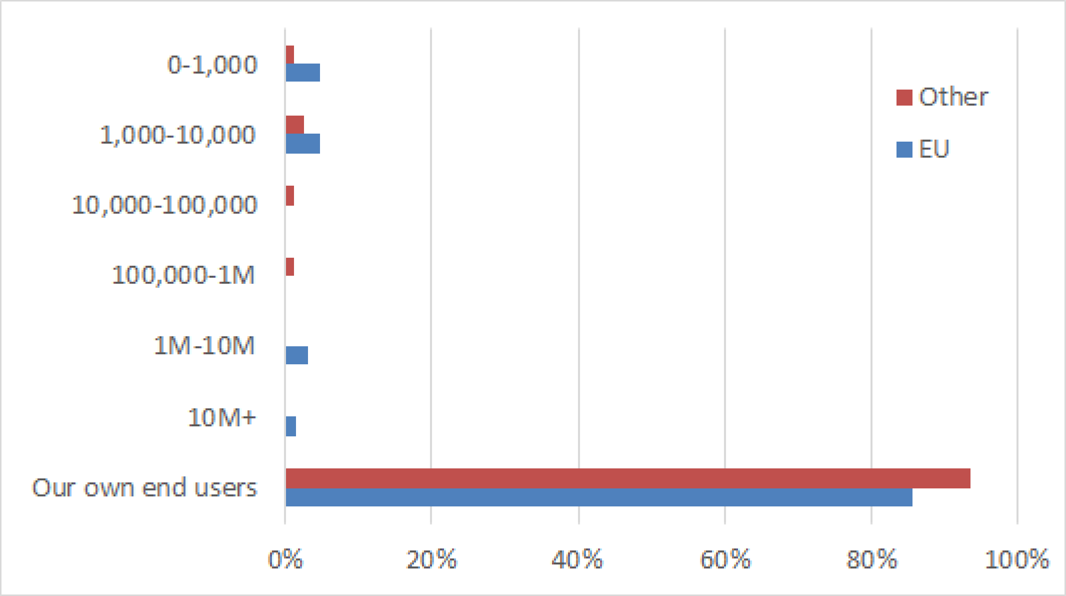


Figure 5: Respondents user type

5.3.5 Question 5

How would you rate the remediation reports with respect to the following?	
Type	Optional
Responses	121 (81%)

Region	**Response**	**Usefulness**	**Timeliness**	**Accuracy**
EU	Very low	0	1	1
	Low	0	1	0
	OK	3	8	8
	Good	32	24	29
	Excellent	24	25	21
	Total	59	59	59
Other	Very low	0	0	0
	Low	2	0	0
	OK	2	4	4
	Good	26	22	26

	Excellent	32	36	32
	Total	62	62	62

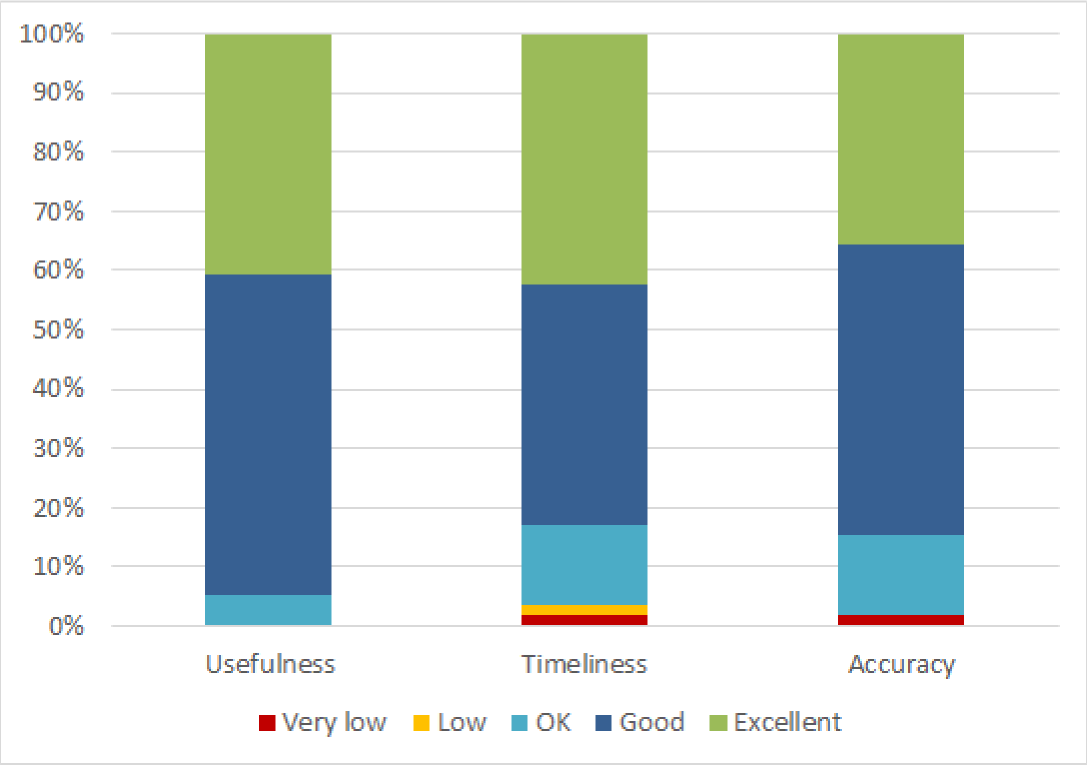


Figure 6: EU-based responses to Question 5

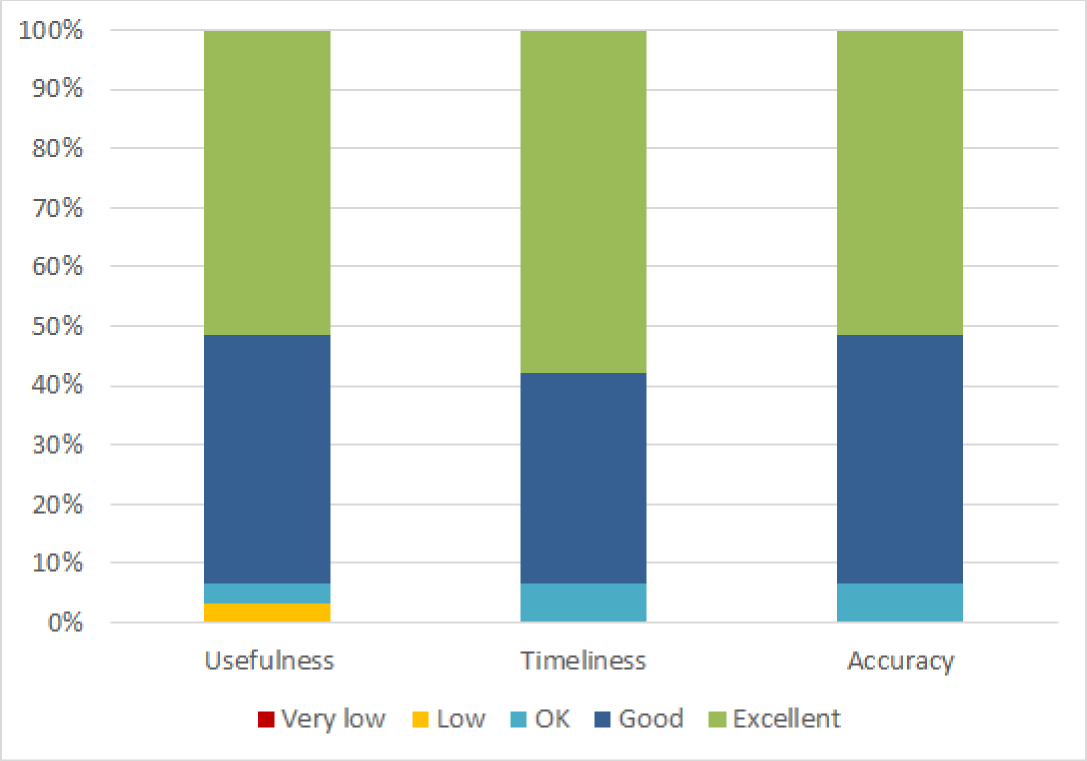


Figure 7: Non-EU based responses to Question 5

5.3.6 Question 6

How useful do you find the following data types delivered by SISSDEN?	
Type	Optional
Responses	96-108 (64-72%)

Region	Response	Brute force attacks	HTTP scanners	ICS scanners	DDoS amplification victims	Darknet traffic
EU	Not useful	0	0	0	1	0
	Somewhat useful	3	3	1	4	3
	Fairly useful	8	13	15	12	14
	Very useful	38	31	27	26	23
	Invaluable	3	8	6	10	7
	Total	52	55	49	53	47
Other	Not useful	0	0	0	1	2
	Somewhat useful	3	4	5	3	3
	Fairly useful	13	7	10	8	7
	Very useful	26	35	29	35	29
	Invaluable	9	6	6	8	8
	Total	51	52	50	55	49

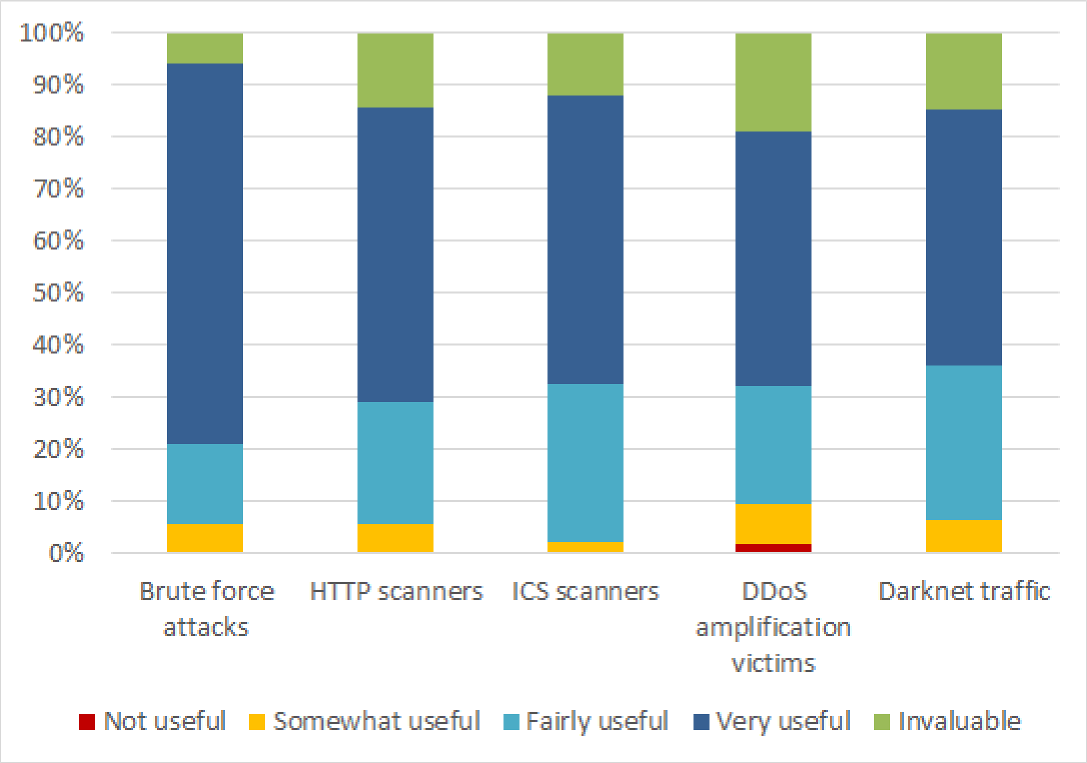


Figure 8: EU-based responses to Question 6

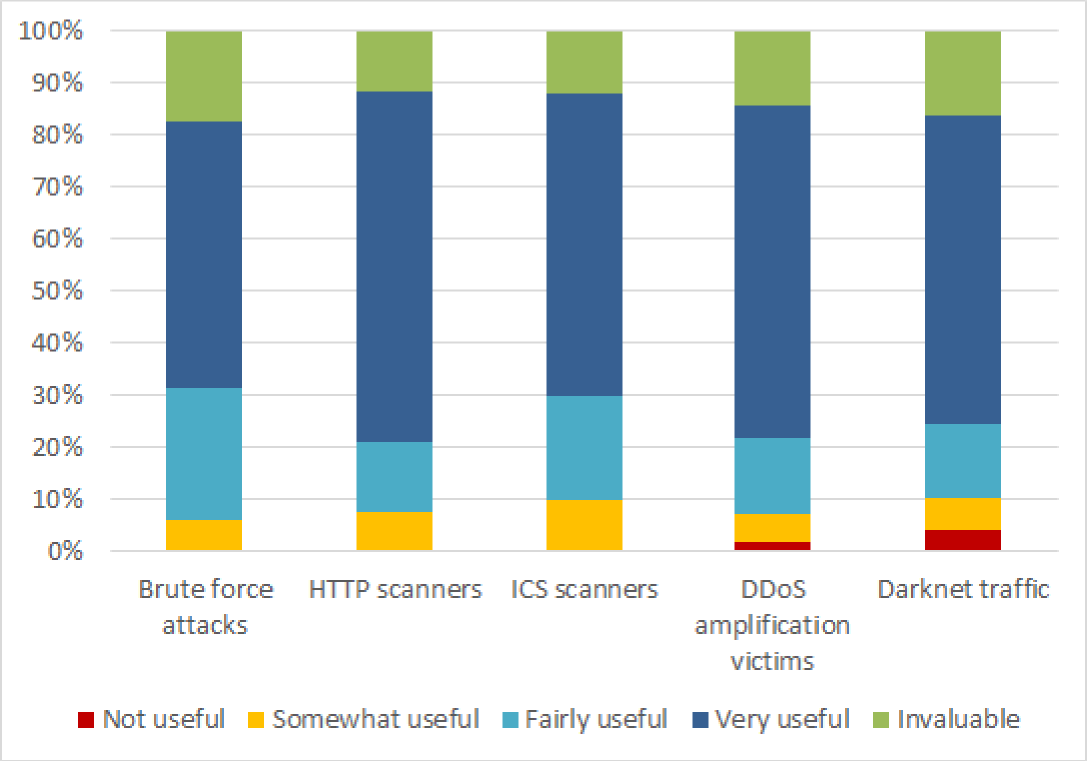


Figure 9: Non-EU based responses to Question 6

5.3.7 Question 7

What proportion of SISSDEN's reports do you think you manage to remediate? Please estimate.	
Type	Optional
Responses	111 (74%)

Region	Response	Count	Percentage
EU	0-10%	4	7.4%
	10-20%	8	14.8%
	20-40%	9	16.7%
	40-60%	6	11.1%
	60-80%	16	29.6%
	80-100%	11	20.4%
	Total	54	100%
Other	0-10%	6	10.5%
	10-20%	7	12.3%
	20-40%	7	12.3%
	40-60%	8	14.0%
	60-80%	11	19.3%
	80-100%	18	31.6%
	Total	57	100%

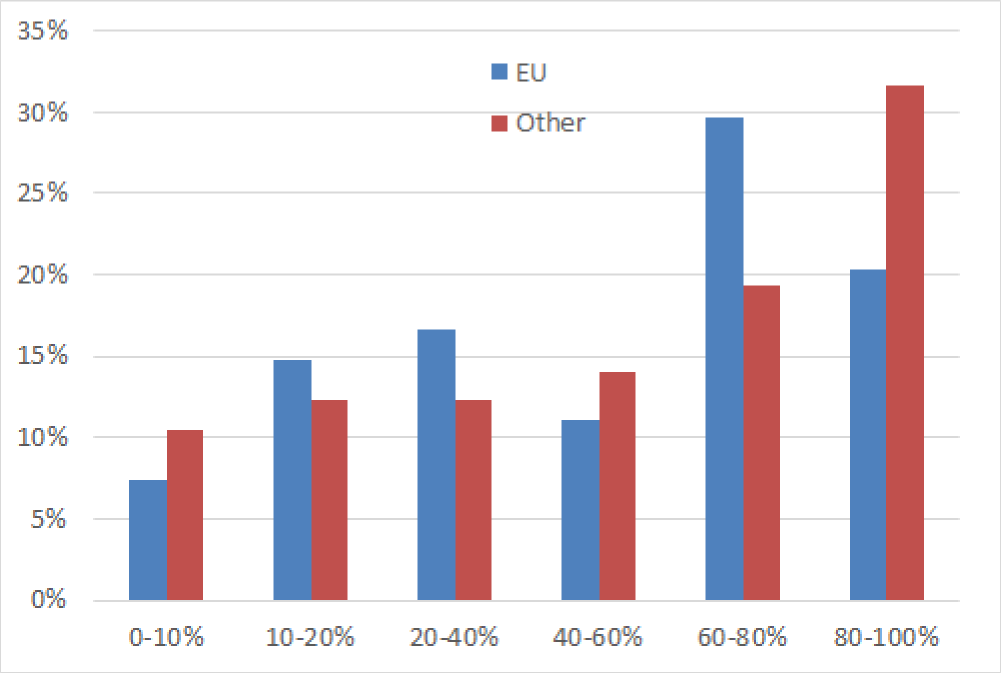


Figure 10: Respondents remediation rate estimation

5.3.8 Question 8

Are there any other data types that you would like SISSDEN to gather or report types you would like to receive?	
Type	Optional free-text field
Responses	23 (15%)

Since this question contained a free-text field, the answers are not published here, since the privacy statement on the survey did not state this. It was considered that if the survey stated that free-text responses may be repeated in a public document, then potential respondents may have been discouraged from participating. The survey was designed to be quick, simple and with mostly optional questions in order to not discourage people from responding.

Instead, the answers have been categorised as shown below. Note that the categories are not distinct - one response may full under multiple categories. Empty or non-actionable responses (e.g. “Not sure”) have been omitted from the categorisation.

Type	**Category**	**Count**
Report types	APT	1
	Compromised sites or machines	2
	Credentials	3

	Criminal campaigns	1
	DoS perpetrators	1
	IoT	2
	Social media	1
	Vulnerabilities	4
Other	Already delivered by SHAD’s reporting	2
	Overall counts or blacklists	3
	Realtime notifications of attacks	1
	Reporting recommendation	1
	Unique responses	**20**

5.3.9 Question 9

Are there any other honeypots or sensors that you would like to see SISSDEN using to collect data? Please enter as many examples as you wish - either specific, or general types.	
Type	Optional free-text field
Responses	10 (7%)

Similarly to Question 8, the answers to this question are categorised below.

Category	**Count**
Already deployed	3
C&C	1
IoT	4
Phishing	1
Tarpits	1
Unique responses	**10**

5.3.10 Question 10

Are you willing to participate in another and more advanced survey?	
Type	Optional
Responses	122 (82%)

Region	Response	Count	Percentage
EU	Yes	34	56.7%
	No	26	43.3%
	Total	60	100%
Other	Yes	36	58.5%
	No	26	41.9%
	Total	62	100%

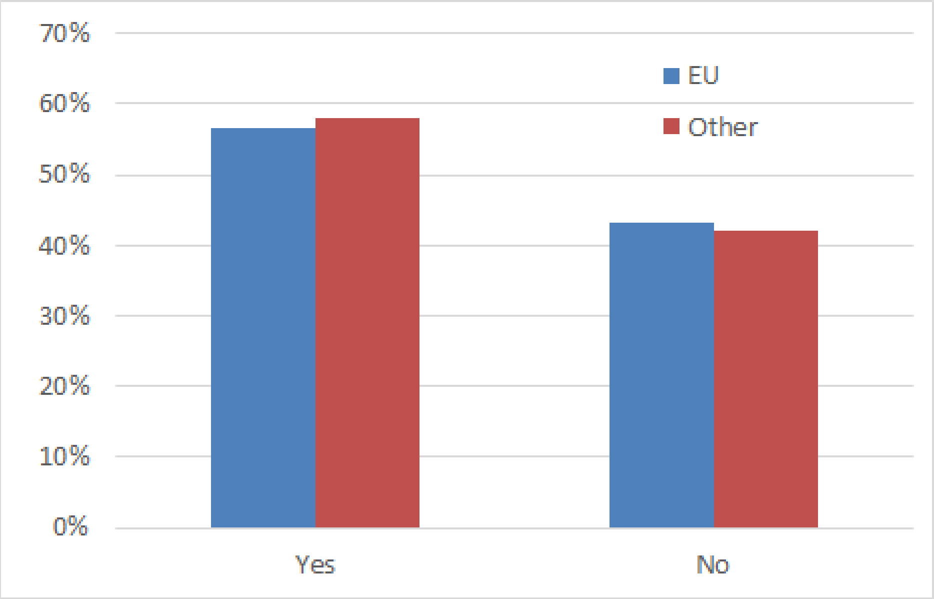


Figure 11: Respondents willing to participate in another survey

5.3.11 Summary

Questions 1 to 4 were aimed at understanding which types of users were responding to the survey. It can be seen that the location of respondents is well distributed across many EU countries, with the addition of a large number of respondents from the United States. In addition, there were respondents from all industry sectors, although the most common was (unsurprisingly) service providers.

Questions 5 and 6 were aimed at understanding how useful the reports are considered to be. Rating the reports in terms of Usefulness, Timeliness and Accuracy, at least 88% of respondents answered a minimum of Good in each category, whilst at least 43% answered Excellent.

Rating the 5 new report types (at the time of the survey), at least 68% of respondents considered all new report types to be Very Useful or better, whilst at least 11% considered them to be Invaluable.

Given the difficulties outlined in *Section 2* in obtaining a quantitative measure of remediation rate, Question 6 was essential in making a qualitative assessment. 77.5% of respondents estimated that at least 20% of SISSDEN's reports were successfully remediated; 63.1 estimated at least 40%; 50.5% estimated at least 60%; and 26.1% estimated at least 80%.

Questions 8 and 9 were aimed at receiving feedback on how SISSDEN can improve its reporting in the future. Notably, several honeypots were recommended which SISSDEN already deploys but does not currently use in remediation reports. This can be due to a variety of factors - for example, not enough data to create a separate report type or data that overlaps with another honeypot.

6 Metrics Dashboard

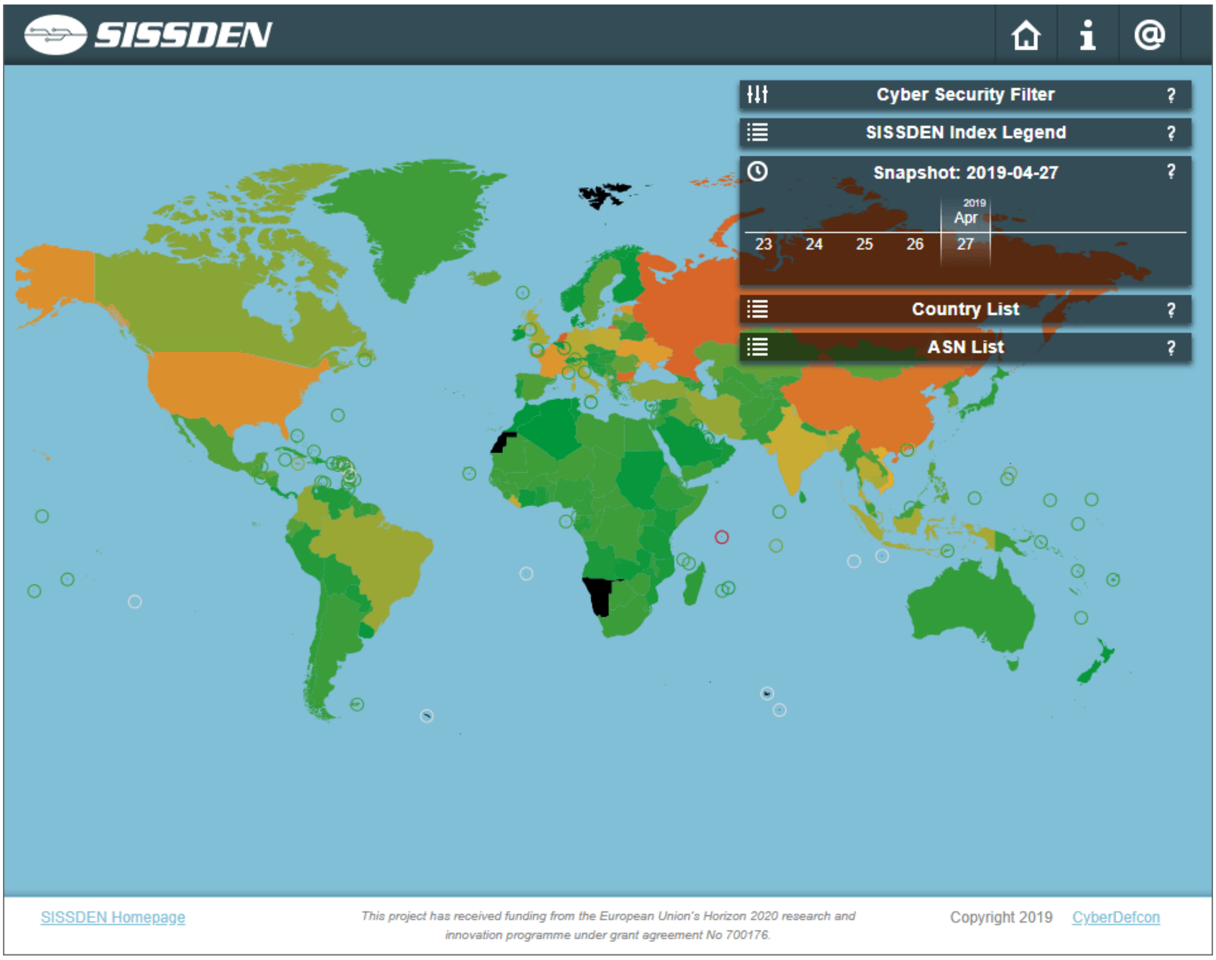
The Metrics Dashboard is published at: https://metrics.sissden.eu

The purpose of the Metrics Dashboard is to display a selection of metrics on a public website to showcase the scale of events that SISSDEN has recorded, and also to provide a means of comparison between the amount of malicious activity served from different regions.

For this reason, it should primarily highlight:

- Region. Being centred around a map view will give it a simple and visual approach.
- Proportional metrics. Since it is easier to compare the values of the proportional metric, this should be the primary metric displayed.

The Metrics Dashboard is based on sideground from CYBE and has been modified with several improvements, and bugs fixed. It has also been refactored to a more generic state to allow the parameters and categories in SISSDEN to be applied.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 700176.

Figure 12: Primary map view

Figure 12 shows the primary map view when the Dashboard is loaded. Countries are colour-coded according to the overall SISSDEN Index for the country.

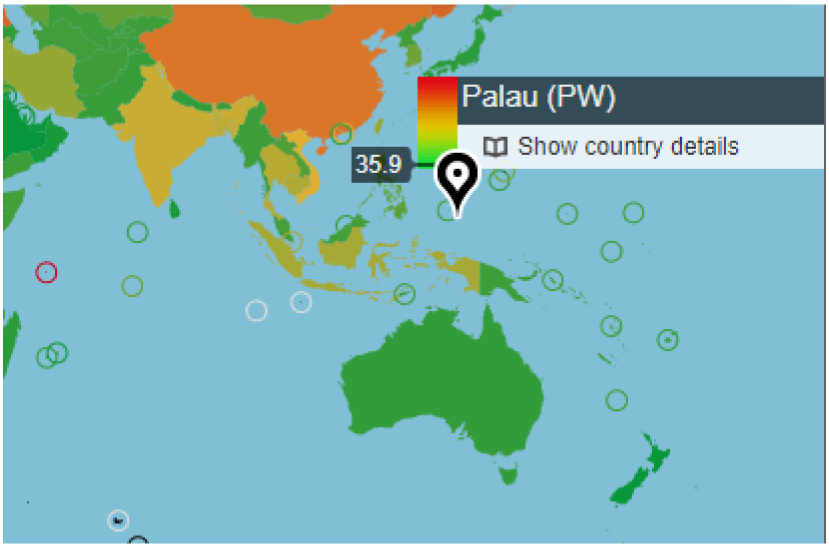
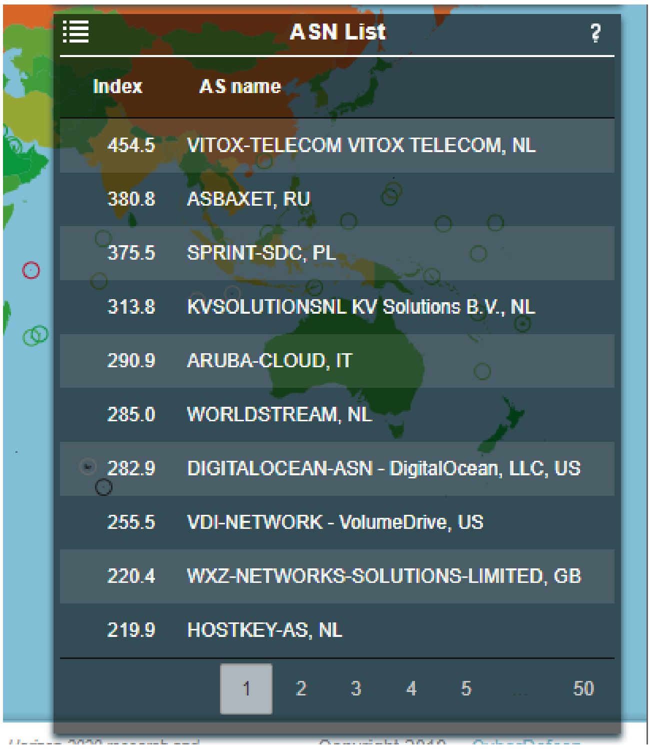


Figure 13: Country legend

Figure 13 shows that smaller countries and islands are highlighted to make it easier to find them. When a country is selected, the Index is displayed for that country along with the country name and a colour scale of the Index.

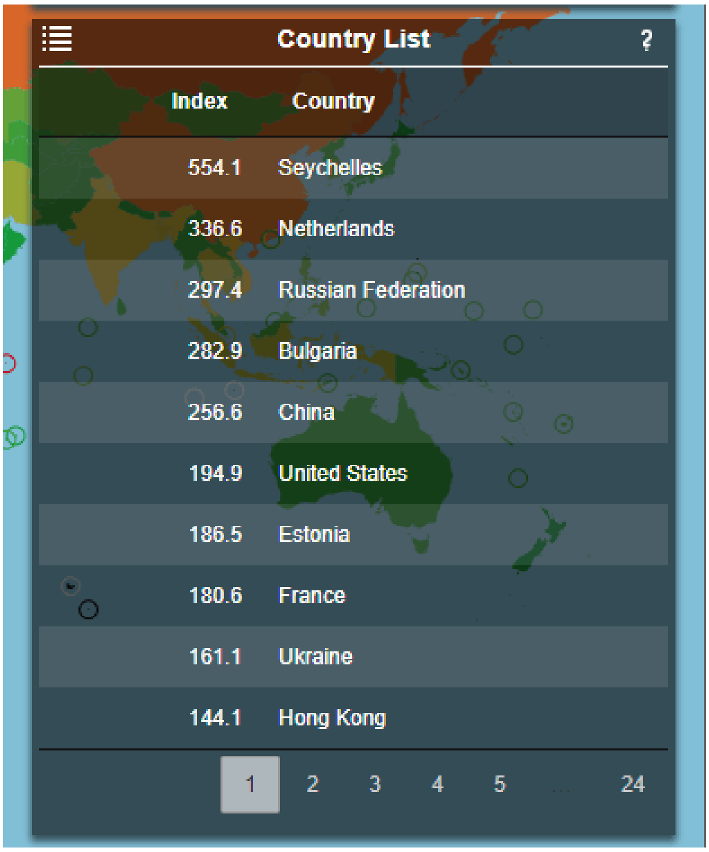


ASN List

Index	AS name
454.5	VITOX-TELECOM VITOX TELECOM, NL
380.8	ASBAXET, RU
375.5	SPRINT-SDC, PL
313.8	KVSOLUTIONSNL KV Solutions B.V., NL
290.9	ARUBA-CLOUD, IT
285.0	WORLDSTREAM, NL
282.9	DIGITALOCEAN-ASN - DigitalOcean, LLC, US
255.5	VDI-NETWORK - VolumeDrive, US
220.4	WXZ-NETWORKS-SOLUTIONS-LIMITED, GB
219.9	HOSTKEY-AS, NL

Figure 14: ASN list

Figure 14 shows the ASN list after being expanded (by being clicked). ASNs are ordered by the SISSDEN Index by default.



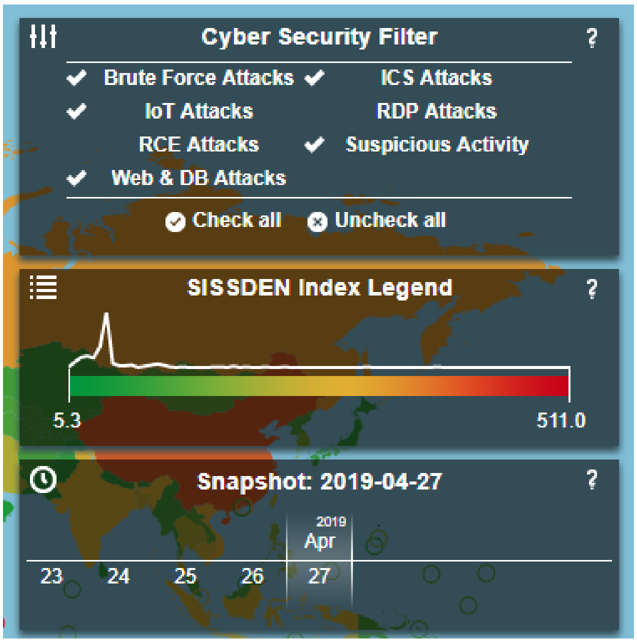
Country List ?

Index	Country
554.1	Seychelles
336.6	Netherlands
297.4	Russian Federation
282.9	Bulgaria
256.6	China
194.9	United States
186.5	Estonia
180.6	France
161.1	Ukraine
144.1	Hong Kong

1 2 3 4 5 ... 24

Figure 15: Country list

Figure 15 shows the country list, which, similarly to the ASN list, is sorted by the Index for the country by default.



Cyber Security Filter ?

✓ Brute Force Attacks ✓ ICS Attacks
✓ IoT Attacks RDP Attacks
RCE Attacks ✓ Suspicious Activity
✓ Web & DB Attacks

✓ Check all ⊗ Uncheck all

SISSDEN Index Legend ?

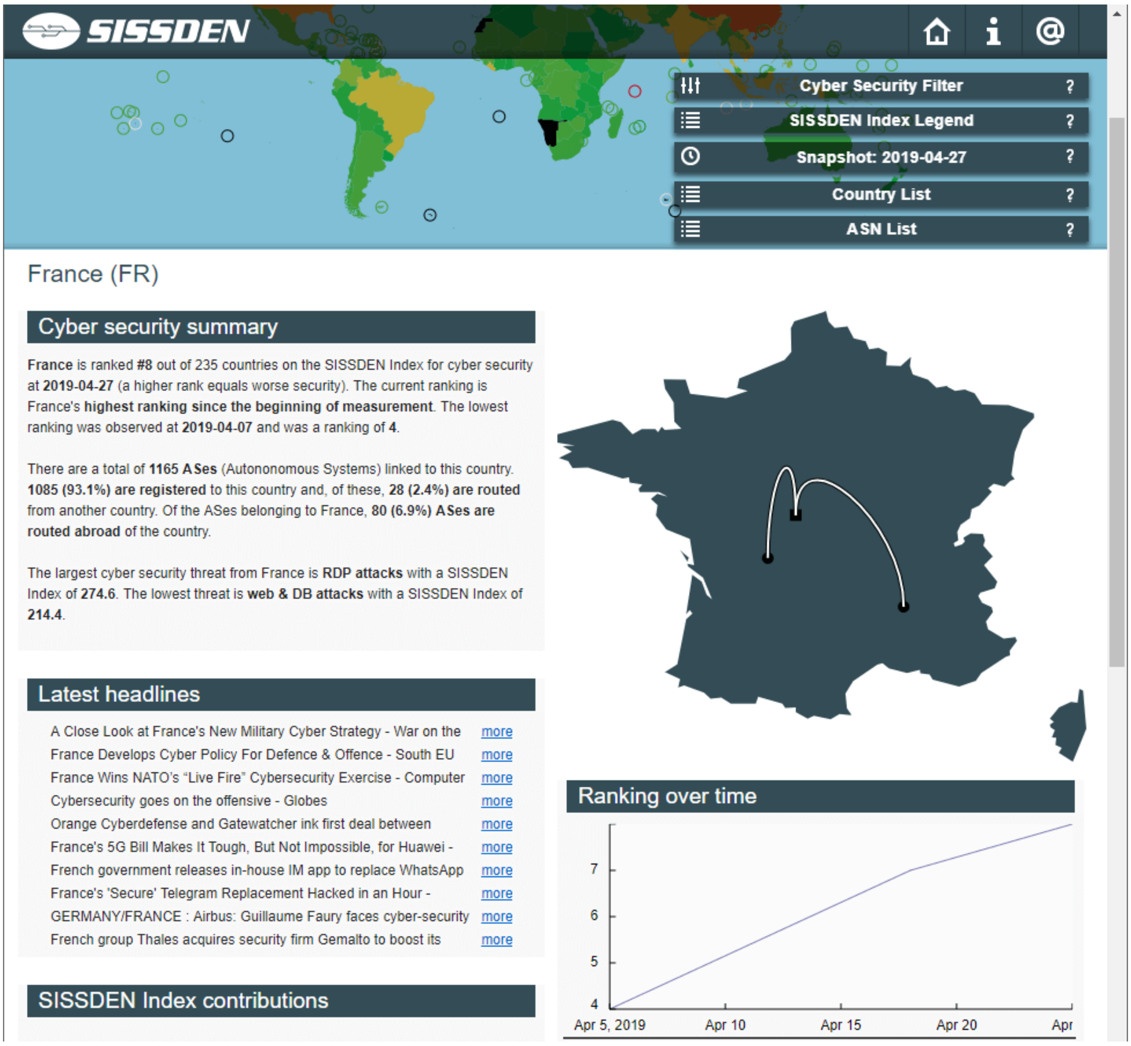
Snapshot: 2019-04-27 ?

Figure 16: Map filters

Figure 16 shows the map controls. The cyber security filter enables categories to be selected and unselected. A custom index is then calculated for the combination of categories and the map automatically updated.

The Index legend displays a colour bar of the Indexes in the current view. The bar is colour coded from green to red, which represent the minimum and maximum Index in the current view. A line plot above the bar shows the distribution of Indexes among all ASNs.

Lastly, a snapshot control enables the metrics to be loaded from a different point in time. When the date is changed, all data is silently reloaded (the map, the Index legend, the country page, etc).



France (FR)

Cyber security summary

France is ranked **#8** out of 235 countries on the SISSDEN Index for cyber security at **2019-04-27** (a higher rank equals worse security). The current ranking is France's **highest ranking since the beginning of measurement**. The lowest ranking was observed at **2019-04-07** and was a ranking of **4**.

There are a total of **1165 ASes** (Autononomous Systems) linked to this country. **1085 (93.1%) are registered** to this country and, of these, **28 (2.4%) are routed** from another country. Of the ASes belonging to France, **80 (6.9%) ASes are routed abroad** of the country.

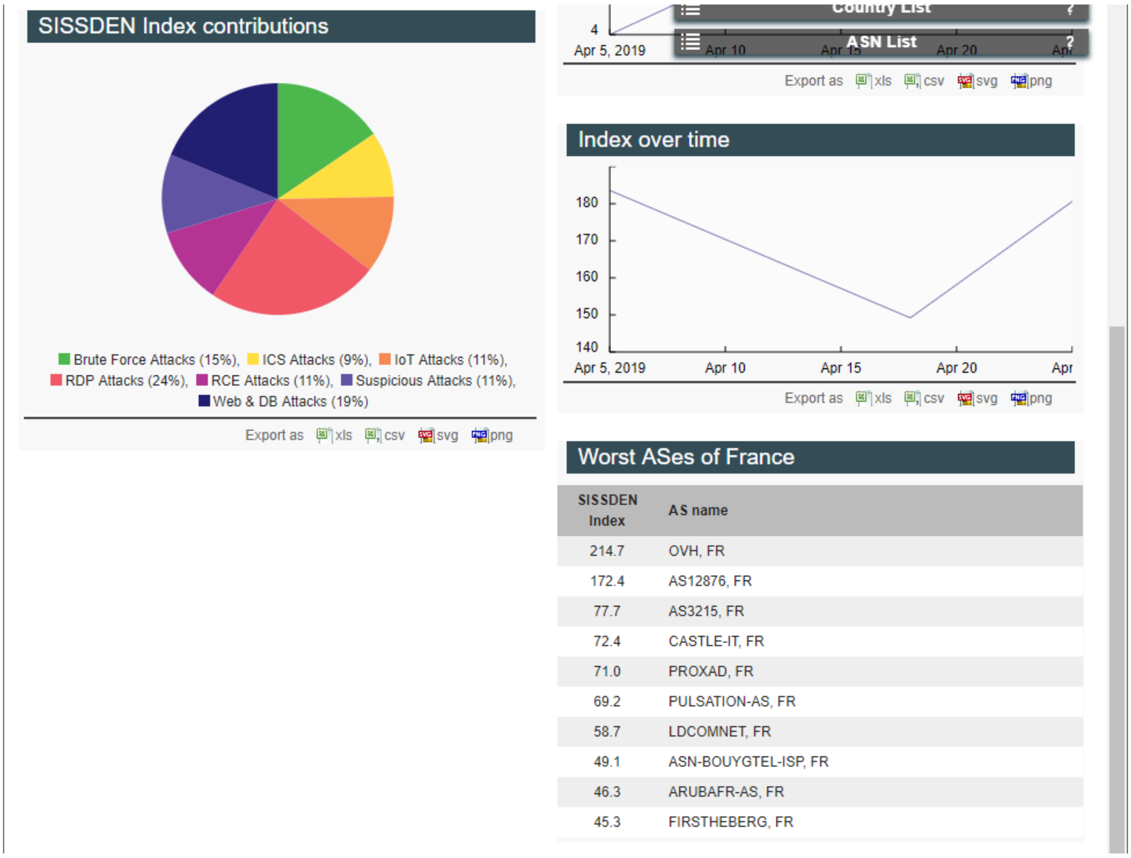
The largest cyber security threat from France is **RDP attacks** with a SISSDEN Index of **274.6**. The lowest threat is **web & DB attacks** with a SISSDEN Index of **214.4**.

Latest headlines

A Close Look at France's New Military Cyber Strategy - War on the	more
France Develops Cyber Policy For Defence & Offence - South EU	more
France Wins NATO's “Live Fire” Cybersecurity Exercise - Computer	more
Cybersecurity goes on the offensive - Globes	more
Orange Cyberdefense and Gatewatcher ink first deal between	more
France's 5G Bill Makes It Tough, But Not Impossible, for Huawei -	more
French government releases in-house IM app to replace WhatsApp	more
France's 'Secure' Telegram Replacement Hacked in an Hour -	more
GERMANY/FRANCE : Airbus: Guillaume Faury faces cyber-security	more
French group Thales acquires security firm Gemalto to boost its	more

SISSDEN Index contributions

Figure 17: Country summary


SISSDEN Index contributions
Brute Force Attacks (15%), ICS Attacks (9%), IoT Attacks (11%),
RDP Attacks (24%), RCE Attacks (11%), Suspicious Attacks (11%),
Web & DB Attacks (19%)
Export as xls csv svg png
Country List
ASN List
4
Apr 5, 2019
Apr 10
Apr 15
Apr 20
Apr
Export as xls csv svg png
Index over time
180
170
160
150
140
Apr 5, 2019
Apr 10
Apr 15
Apr 20
Apr
Export as xls csv svg png

Worst ASes of France

SISSDEN Index	AS name
214.7	OVH, FR
172.4	AS12876, FR
77.7	AS3215, FR
72.4	CASTLE-IT, FR
71.0	PROXAD, FR
69.2	PULSATION-AS, FR
58.7	LDCOMNET, FR
49.1	ASN-BOUYGTEL-ISP, FR
46.3	ARUBAFR-AS, FR
45.3	FIRSTHEBERG, FR

Figure 18: Country summary (cont'd)

Figure 17 and *Figure 18* display the country details page. A text summary describes the main statistics, and a larger image of the country is displayed, with Internet Exchange locations highlighted. There are also charts of the ranking and Index over time, a breakdown of the category contributions (to view the worst/best categories for that country at a glance) and a list of the worst ASNs in the country. When the snapshot is updated, this page also updates.

6.1 Post-project

The Metrics Dashboard will be maintained beyond the project for 3 years. This will include iterative development to ensure that the data can continue to be accessed in ways that users request. For example, if there is a user demand for access via daily email feeds, PDFs, etc then this will be provided.

7 Conclusions

The SISSDEN project provided a large-scale data collection from honeypots and darknets that was the ideal basis for metrics and further research, including via the Curated Reference Data Set. This document presents a set of metrics applied to this data set, comprised of absolute and relative indexes across 7 different threat categories.

Whilst all of SISSDEN’s data was not used as input for the metrics, instead a large subset of the data was selected which enabled a single metric to be compared between different categories. This made the metric ideal for publishing on the Metrics Dashboard component where the public can view and compare the indexes between different countries, ASNs and time periods, and where it will be maintained beyond the project.

The results of the proportional metric were normalised by IP address, meaning that some smaller countries and ASNs with relatively large amounts of attacks detected by SISSDEN were ranked highly by this metric. This produced some interesting results, such as Seychelles being listed on the dashboard as the worst-ranked country.

Despite the large-scale data collection achieved within SISSDEN, the reporting system is effectively unidirectional – victim reports are sent out on a daily basis, but there exists no official mechanism for receiving information back on these reports on a large scale. In order to effectively measure remediation rates at scale, significant innovation would be required, not only to the reporting system but also to the internal processes of reporting recipients, many of which do not currently store the information that would be required to accurately assess which incidents were remediated due to SISSDEN. Further research is required here in order to provide not only SISSDEN but other initiatives with the necessary information to be able to measure the effectiveness of their remediation efforts.

In order to assess the effectiveness of SISSDEN’s reporting specifically, a survey was carried out which showed overwhelmingly that recipients found SISSDEN’s reporting to be useful, timely and accurate. The 5 new reporting types introduced by SISSDEN were also positively received.