MANANA: motivating questions

Have you ever wondered...

- **who** *your mobile apps are communicating with*, in their everyday functioning? and **how much**?

- do these communication happen in **clear** or **encrypted**?

- where are located the **servers** that allow the app to function?

- which **countries** are **traversed** by your traffic?

- which countries the **managers of the network** infrastructure are accountable to?

- is this **behavior** the same for different apps that provide the **same kind of service**?

NET MAKERS

MANANA: impact

Besides curiosity (personal, or research-driven)

- the global communication infrastructure has been increasingly subject of **cyberattacks**

- **national governments** have been **regulating mobile apps usage** and enforcing Internet censorship by different means, including **disrupting Internet connectivity**

- **functionally-similar apps** can have a significantly **different impact (and dependence)** on the global communication infrastructure

> with no way of observing the geopolitic network footprint, **the user cannot enact informed choices**

https://freedomhouse.org/report/freedom-net/2024/struggle-trust-online

MANANA: goals

> with no way of observing the geopolitic network footprint, **the user cannot enact informed choices**

To help in this matter, MANANA provides for monitored apps

- easy-to-read **characterization of network traffic**
  - received/sent, bytes/connections, encrypted/clear
- **countries** traversed or reached
- countries of the **network device manager** (WHOIS)
- a summarizing index of **"information freedom risk"** associated with the country, weighted according to traffic

WORK SHOP GARR 2024   NET MAKERS

MANANA: design principles

> with no way of observing the geopolitic network footprint, **the user cannot enact informed choices**

to this aim, MANANA followed these principles

- ease of use (**laymen oriented**)
- **no rooting** of the device is required
- user **privacy** must be guaranteed
- **incentives** to usage and support of the community of users
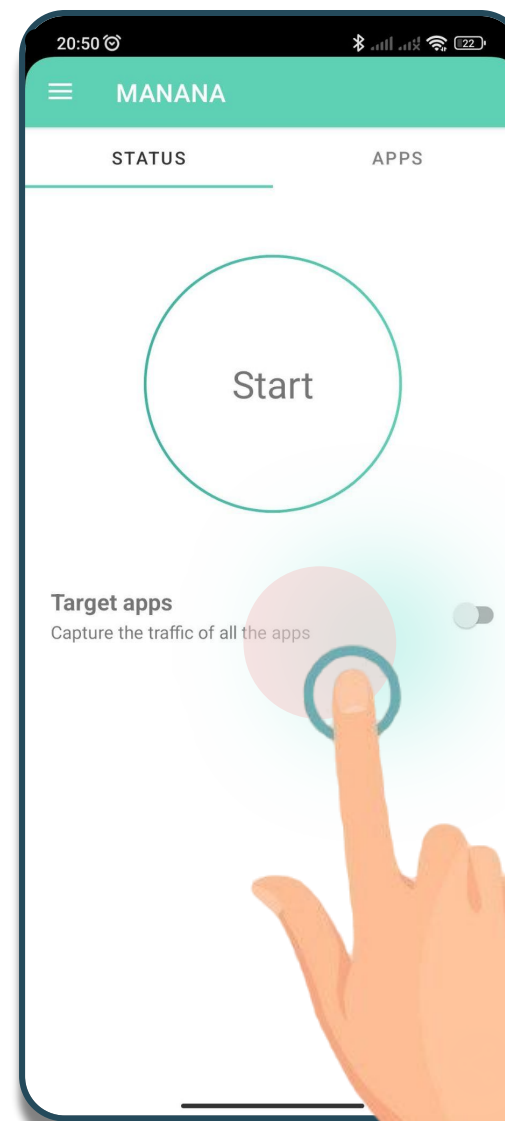- gather **crowdsourced data** for scientific studies

MANANA: use

(1) **launch MANANA**

(2) select the target apps

- the selection is kept over restarts

(3) start the traffic observation

(4) inspect the results

(5) share the results with the community

(6) REPEAT!

WORK SHOP GARR 2024

NET MAKERS
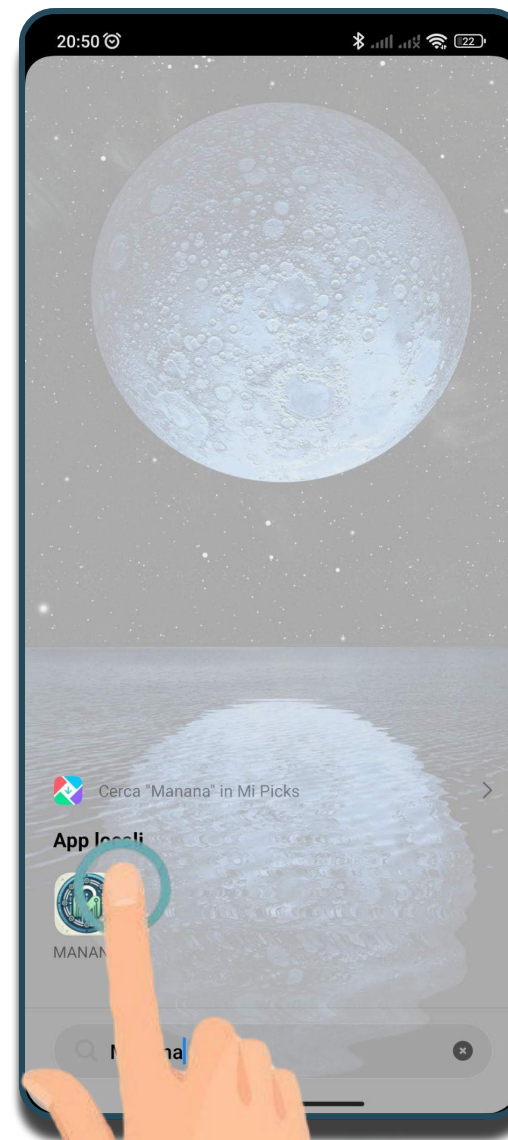
MANANA: use

(1) launch MANANA

(2) **select the target apps**
   - the selection is kept over restarts

(3) start the traffic observation

(4) inspect the results

(5) share the results with the community

(6) REPEAT!

MANANA: use

(1) launch MANANA

(2) select the target apps
- the selection is kept over restarts

(3) **start the traffic observation**

(4) inspect the results
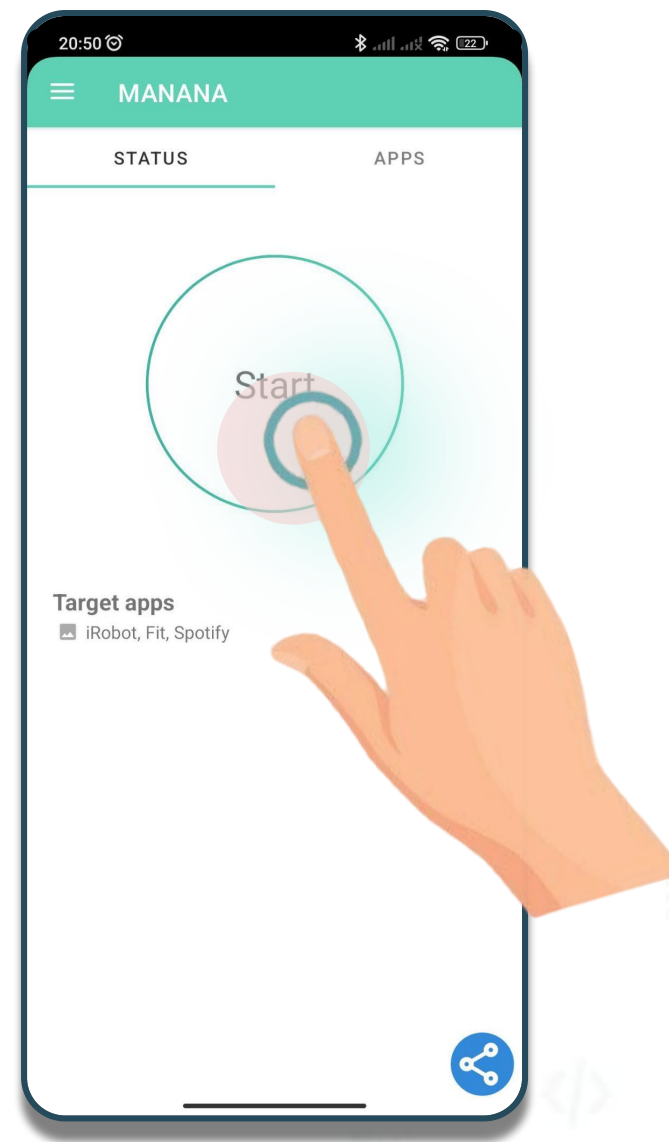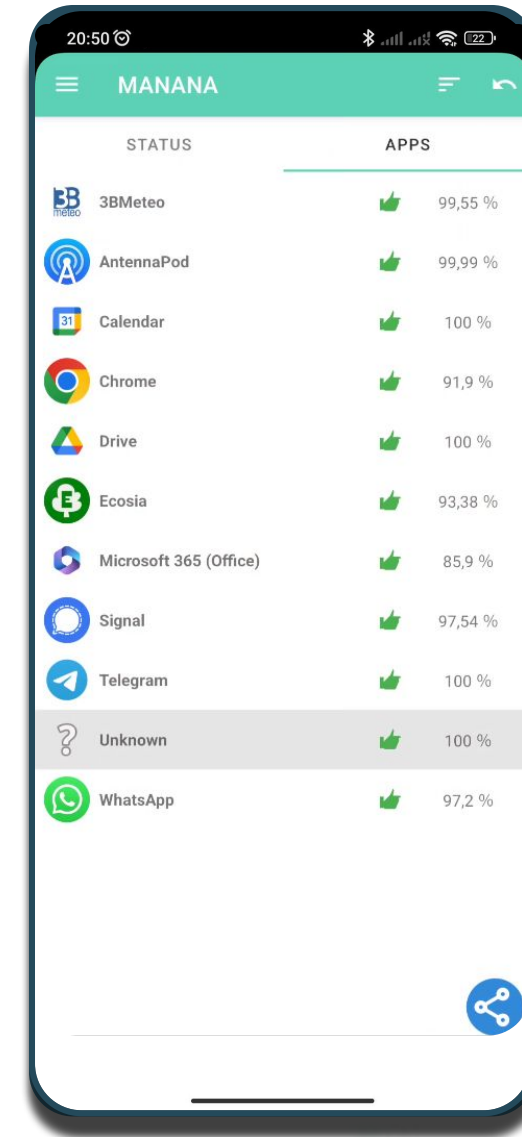
(5) share the results with the community

(6) REPEAT!

Overall information safety report

(1) launch MANANA

(2) select the target apps

    - the selection is kept over restarts

(3) start the traffic observation

(4) **inspect the results**

(5) share the results with the community

(6) REPEAT!

Overall information safety report

(1) launch MANANA

(2) select the target apps
  - the selection is kept over restarts

(3) start the traffic observation

(4) **inspect the results**

(5) share the results with the community

(6) REPEAT!

a single summarizing value (**information safety index***) is provided for each app
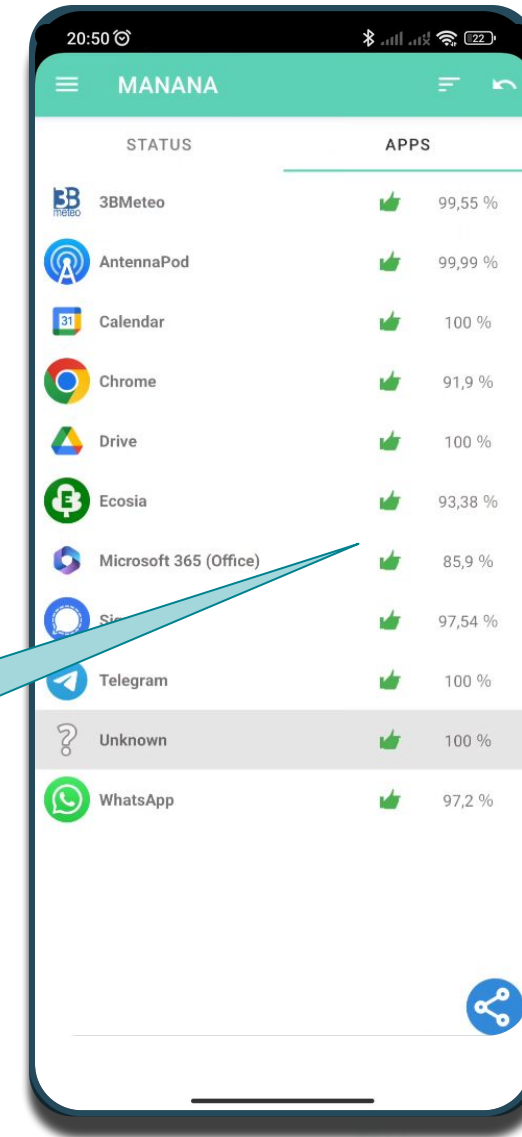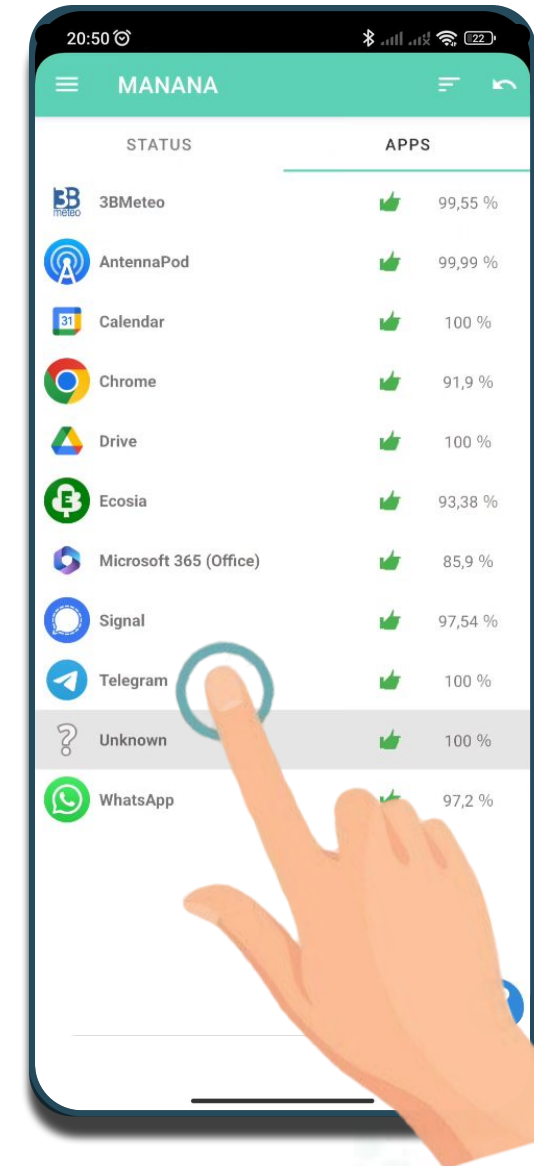
* more details later

Overall information safety report

(1) launch MANANA

(2) select the target apps

- the selection is kept over restarts

(3) start the traffic observation

(4) **inspect the results**

(5) share the results with the community

(6) REPEAT!

for **each app**, a traffic profile is reported, in the form of **"nutrition labels"**

MANANA: information safety labels

(1) launch MANANA

(2) select the target apps
  - the selection is kept over restarts

(3) start the traffic observation

(4) **inspect the results**
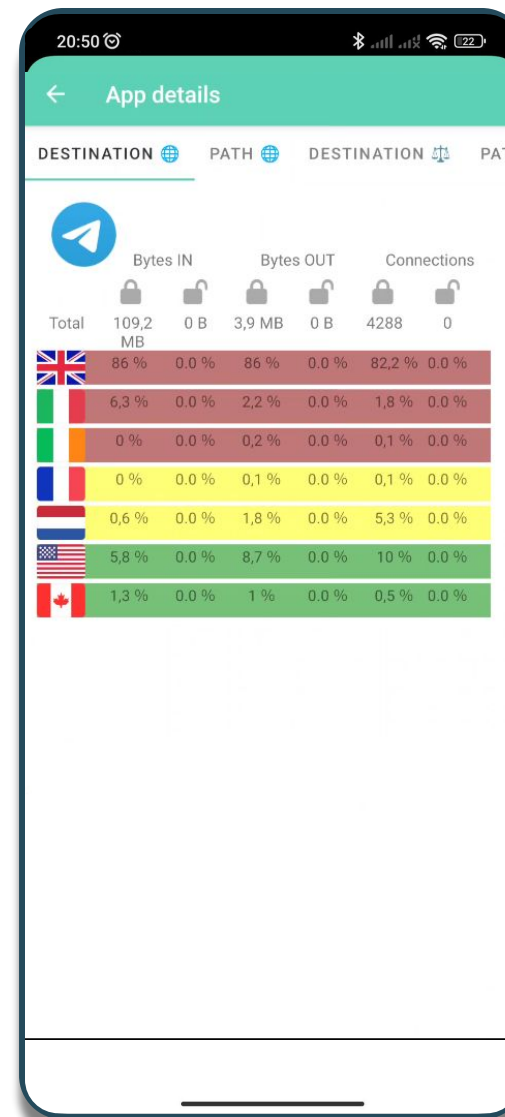
(5) share the results with the community

(6) REPEAT!

for **each app**, a traffic profile is reported, in the form of **"nutrition labels"**

DESTINATION 🌐 GEOGRAPHIC COMPOSITION

on columns get the breakdown of traffic

Bytes IN/OUT, connections

further divided in
🔒 encrypted
🔓 not encrypted

in total for the app,
and by **server country**
(**destination IP**, geodb)

this represents physical dependence/exposition
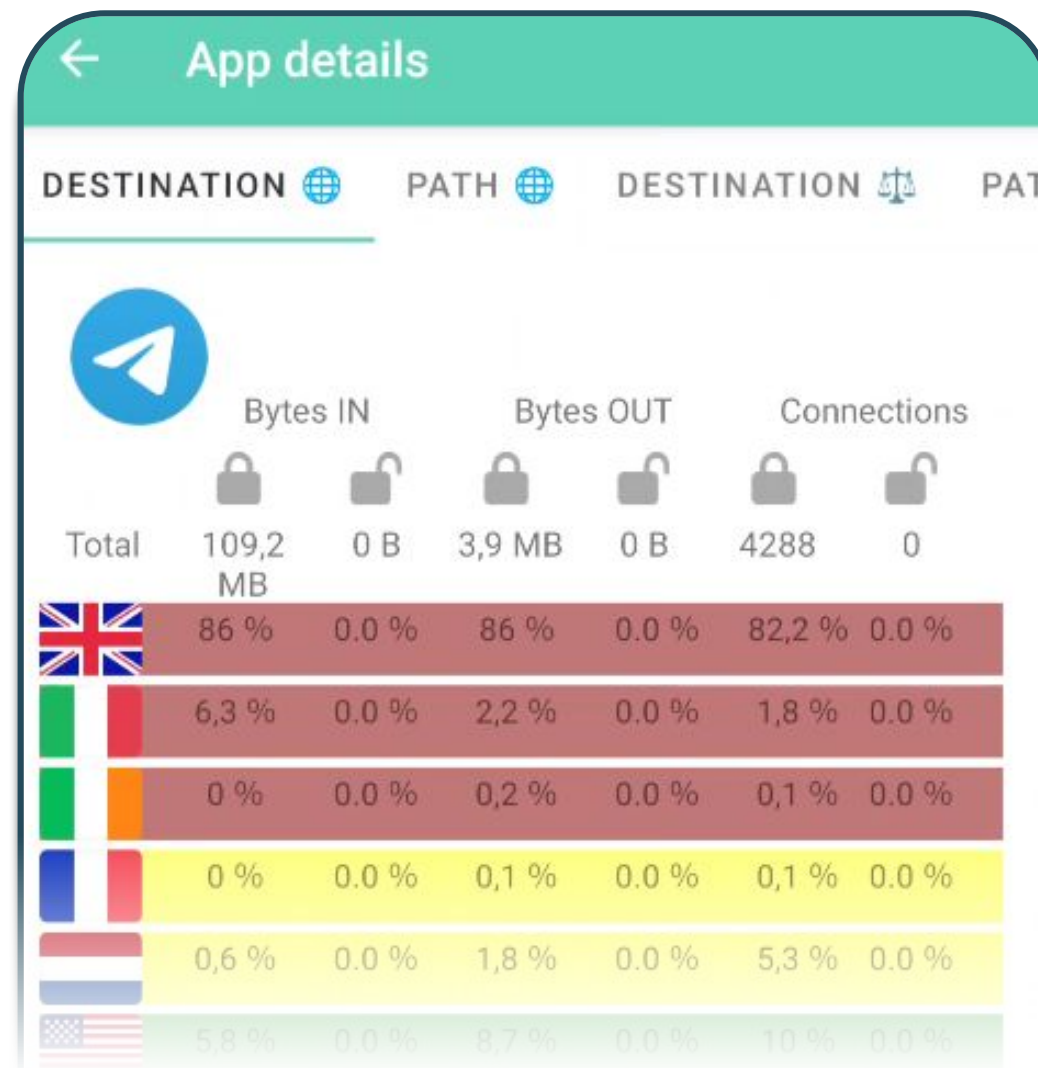
PATH 🌐
GEOGRAPHIC COMPOSITION

Same kind of information, for **traversed countries** (**intermediate traceroute IPs**, geodb).

In this case the percentage shows **which share passed through a given country**.

Only "passing through" countries are considered: neither the source (100%), nor the destination (shown in previous tab).
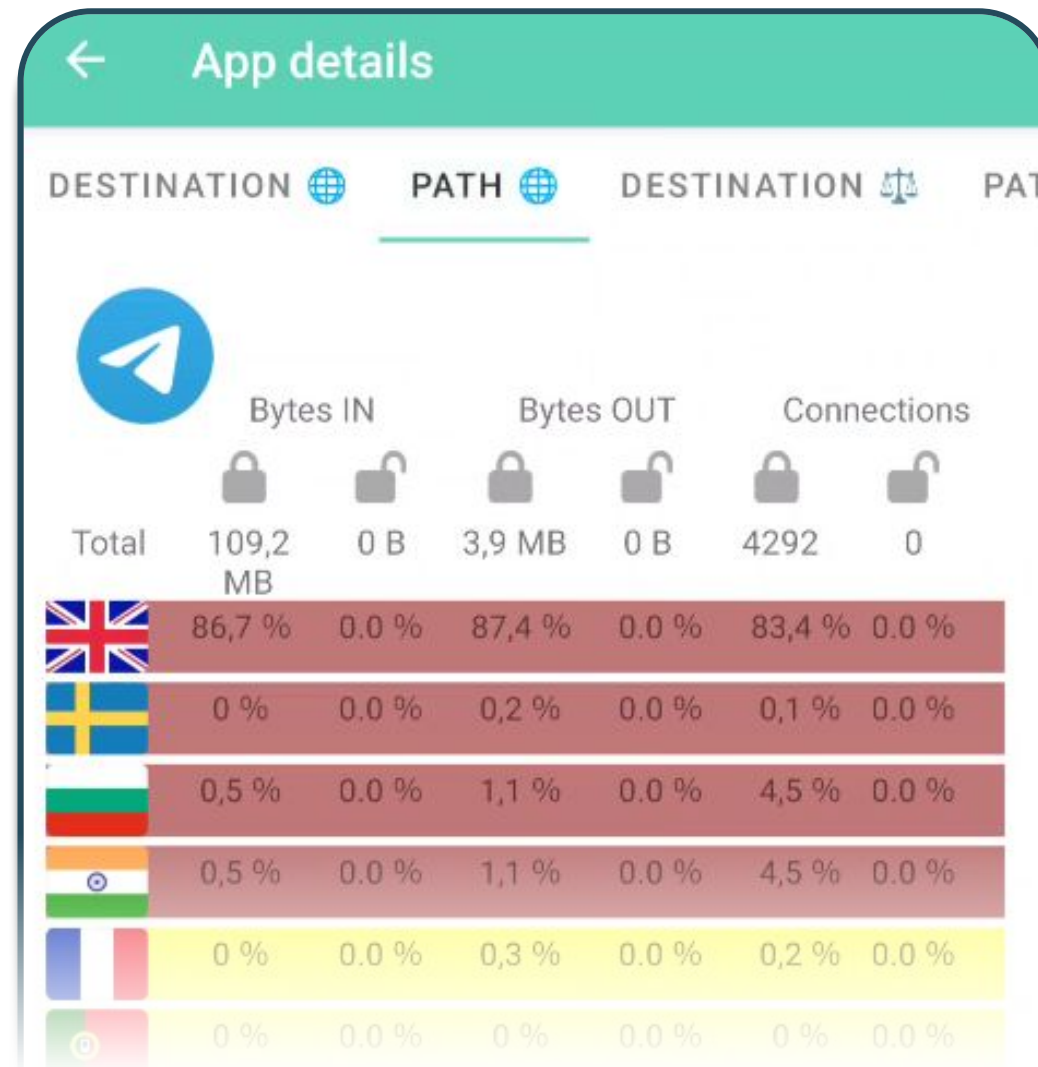
DESTINATION ADMINISTRATIVE COMPOSITION

Same kind of information, for **administrative destinations** (**destination IPs**, WHOIS).

In this case the percentage shows which share is sent to **hosting servers** whose **manager is legally located** in a given country

(exposed to lawful interception, or different degrees of censorship)
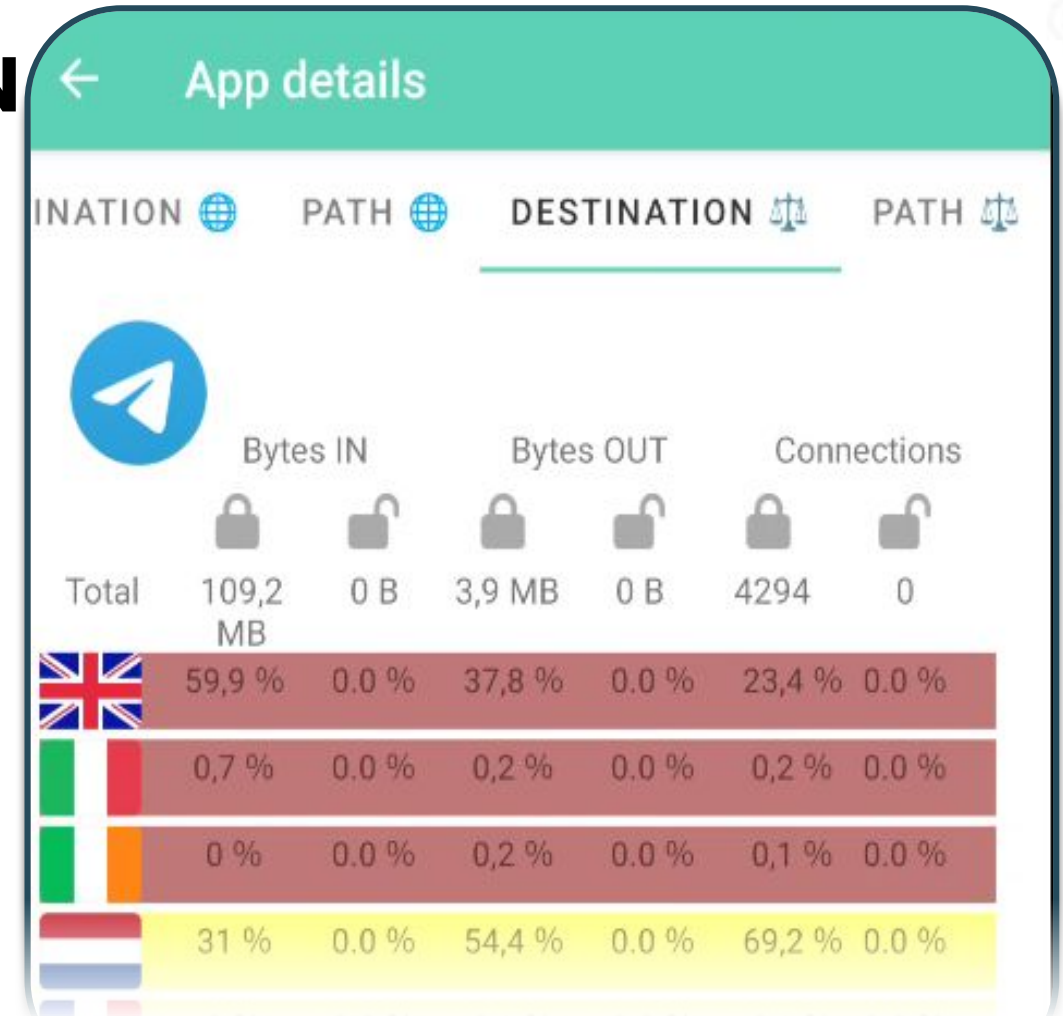
PATH ⚖
ADMINISTRATIVE COMPOSITION

Same kind of information, for **administrative destinations** (**intermediate traceroute IPs**, WHOIS).

In this case the percentage shows which share passed **through a network device** whose **manager is legally located** in a given country.

(exposed to lawful interception, or different degrees of censorship)
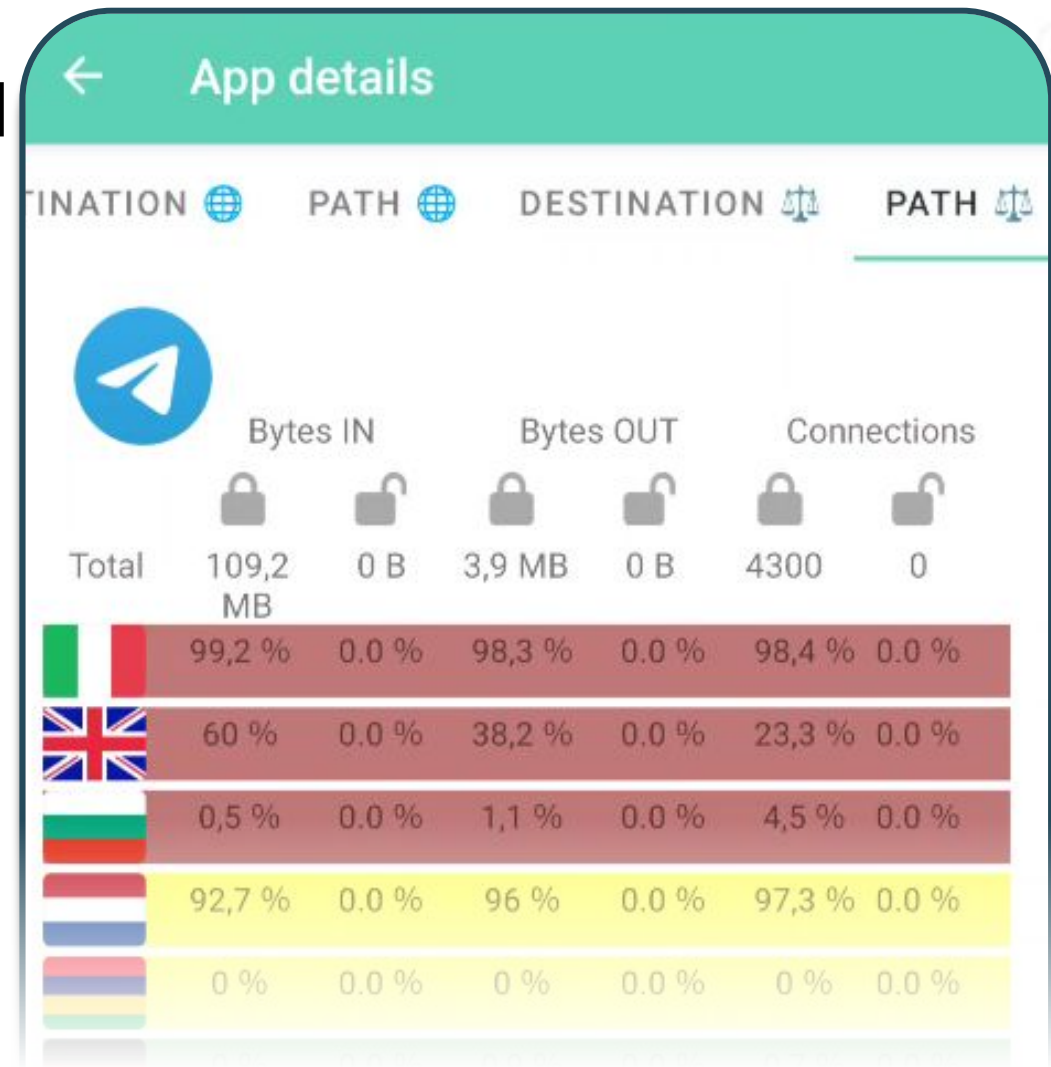
COUNTRY DETAIL AND CLASSIFICATION

Clicking on a country flag, the relative details are shown

Besides traffic sent/received the **country type** is shown.

This is derived from **V-DEM** project* dataset, specifically values about
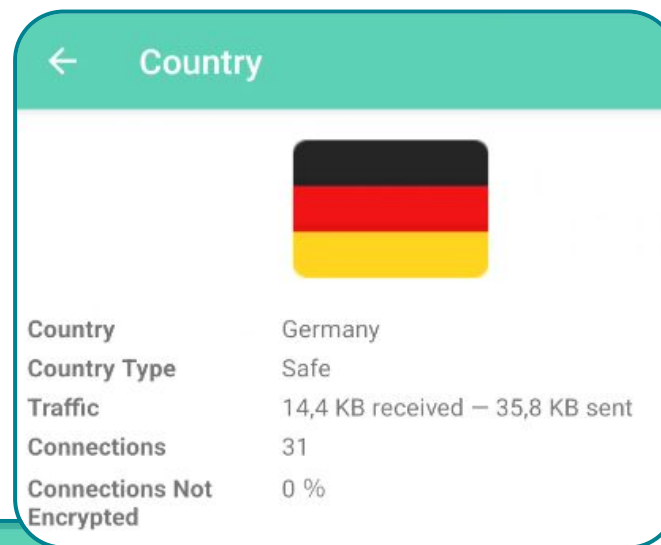
- <u>Internet censorship</u> effort
- Government <u>Internet filtering</u> in practice
- Government <u>Internet shut down</u> capacity
- Government <u>Internet shut down</u> in practice
- <u>Privacy protection</u> by law content

the **minimum** (worst) value among the listed fields is mapped as follows**

- >=3 :   *Safe*
- in]1,3[:   *Almost Safe*
- <=1 :   *Not Safe*

** this is arbitrarily chosen as a first working hypothesis, to be refined.
* by V-Dem Institute, Dept. of Political Science, University of Gothenburg, Sweden. https://www.v-dem.net/

MANANA: use case

(1) launch MANANA

(2) select the target apps

- the selection is kept over restarts

(3) start the traffic observation

(4) inspect the results

(5) **share the results with the community**

(6) REPEAT!

MANANA: use case

(1) launch MANANA

(2) select the target apps

- the selection is kept over restarts

(3) start the traffic observation

(4) inspect the results

(5) **share the results with the community**

(6) REPEAT!

MANANA: use case

(1) launch MANANA

(2) select the target apps
  - the selection is kept over restarts

(3) start the traffic observation

(4) inspect the results

(5) share the results with the community

(6) **REPEAT**!

WORK SHOP GARR 2024   NET MAKERS

Uploaded files: connections report .csv

This file contains data regarding the various observed connections. Source IP is not logged. Each connection (5-tuple) reports the following values:

- **IPProto**: Protocol field of the IP Packet
- **SrcIP**: Source IP (<u>private: local VPN</u>)
- **SrcPort**: Source port number
- **DstIp**: Destination IP
- **DstPort**: Destination Port
- **UID**: App identificator
- **App**: App Name
- **Country**: Destination IP <u>geolocation</u>
- **ASN**: Destination IP related ASN

- **ClassificationValue**: Country information-safety Classification
- **Proto**: Application Level Protocol (from **ntop** classification library)
- **Status**: Connection Status
- **Info**: domain (from observed DNS queries)
- **BytesSent**, **BytesRcvd**
- **PktsSent**, **PktsRcvd**
- **FirstSeen**: Timestamp* First Seen
- **LastSeen**: Timestamp* Last Seen

* **timestamps are anonymized** with **Random Shifts** (anticipated up to 2h) - see <u>Sec.4.3 RFC6235</u>. A different offset is drawn for each observation session.

WORK SHOP GARR 2024   NET MAKERS

report .csv example

... plus first/last packet timestamps

IPProto	SrcIP	SrcPort	DstIp	DstPort	UID	App	Country	ASN	ficatior	Proto	Status	Info	BytesSent	BytesRcvd	PktsSent	PktsRcvd
17	10.215.173.1	27736	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	mdm2.asus.com	59	59	1	1
17	10.215.173.1	16359	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	mdm2.asus.com	59	59	1	1
17	10.215.173.1	20869	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	portal.fb.com	59	118	1	1
6	10.215.173.1	60834	157.240.231.15	80	10071	Facebook	Italy	AS32934 - Facebook, Inc.	1	HTTP	Active	portal.fb.com	357	284	4	3
6	10.215.173.1	60834	157.240.231.15	80	10071	Facebook	Italy	AS32934 - Facebook, Inc.	1	HTTP	Active	portal.fb.com	0	0	4	3
17	10.215.173.1	20658	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	whois.lacnic.net	62	78	1	1
17	10.215.173.1	32361	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	clients3.google.com	65	115	1	1
17	10.215.173.1	14543	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	static.whatsapp.net	65	118	1	1
6	10.215.173.1	35442	157.240.231.60	443	10140	WhatsApp	Italy	AS32934 - Facebook, Inc.	1	HTTPS	Active	static.whatsapp.net	912	3698	9	8
6	10.215.173.1	35442	157.240.231.60	443	10140	WhatsApp	Italy	AS32934 - Facebook, Inc.	1	HTTPS	Active	static.whatsapp.net	0	0	9	8
6	10.215.173.1	59725	216.58.204.142	80	10140	WhatsApp	d Kingdom	AS15169 - Google LLC	1	HTTP	Active	clients3.google.com	359	255	4	3
6	10.215.173.1	59725	216.58.204.142	80	10140	WhatsApp	d Kingdom	AS15169 - Google LLC	1	HTTP	Active	clients3.google.com	0	0	4	3
17	10.215.173.1	8624	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	asia.pool.ntp.org	63	127	1	1
17	10.215.173.1	53746	162.159.200.1	123	10140	WhatsApp	Canada	AS13335 - Cloudflare, Inc.	3	NTP	Active	asia.pool.ntp.org	76	76	1	1
17	10.215.173.1	53746	162.159.200.1	123	10140	WhatsApp	Canada	AS13335 - Cloudflare, Inc.	3	NTP	Active	asia.pool.ntp.org	0	0	1	1
17	10.215.173.1	19303	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	whois.ripe.net	60	76	1	1
17	10.215.173.1	5242	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	whois.arin.net	60	108	1	1
17	10.215.173.1	24242	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	whois.apnic.net	61	77	1	1
17	10.215.173.1	26689	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	whois.nic.or.kr	61	77	1	1
17	10.215.173.1	10349	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Closed	mtalk.google.com	62	117	1	1
6	10.215.173.1	47527	108.177.96.188	5228	10029	ip Transport	ited States	AS15169 - Google LLC	3	TLS	Active	mtalk.google.com	1286	1184	8	8
6	10.215.173.1	47527	108.177.96.188	5228	10029	ip Transport	ited States	AS15169 - Google LLC	3	TLS	Active	mtalk.google.com	0	0	8	8
17	10.215.173.1	19143	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Active	android.apis.google.com	69	0	1	0
17	10.215.173.1	19143	10.215.173.2	53	0	Root	No Info	Unknown ASN	-1	DNS	Active	android.apis.google.com	0	119	1	0
6	10.215.173.1	54438	142.251.209.14	443	10029	ip Transport	Italy	AS15169 - Google LLC	1	HTTPS	Active	android.apis.google.com	1492	1683	7	6
6	10.215.173.1	54438	142.251.209.14	443	10029	ip Transport	Italy	AS15169 - Google LLC	1	HTTPS	Active	android.apis.google.com	0	0	7	6

WORK SHOP GARR 2024

NET MAKERS

Uploaded files: traceroute report .json (1/2)

This file reports data regarding the <u>traceroute result</u> (for every connection **DstIP**)...

Every app is characterised by (1):

- **AppName**
- **traceCountries:** a list of IPs, associated to the app, each with following data.
  - the **result** of the Traceroute towards that IP
  - **timeStamp** at the end of the Traceroute
  - **orderedListCountry,** the countries associated to the IPs found from the Traceroute and <u>geolocated</u>
  - **sentBytesEnc**, encrypted, **sentBytesNotEnc**, unencrypted
  - **rcvdBytesEnc**, **rcvdBytesNotEnc**
  - **numConnsEnc**, **numConnsNotEnc.**
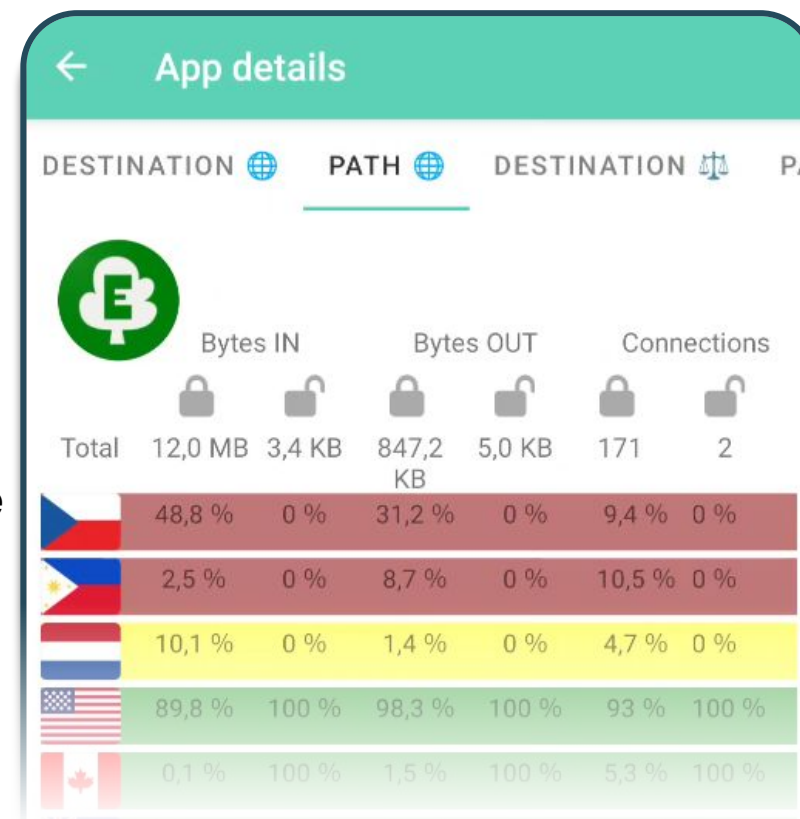
Uploaded files: traceroute report .json (2/2)

This file reports also data regarding <u>Whois query</u>
(for every connection **DstIP and IPs** of the trace).

Every app is characterised also by (2):
- **WhoisTraceCountries,** same format as traceCountries but the **orderedListCountry** from whois queries

- **WhoisDestinationCoutries,** a list of countries, each with:
  - **IpList:** the IPs associated to that country and the Timestamp associated to the Whois query
  - **classificationValue** associated to the country
  - **rcvdBytesEnc**, **rcvdBytesNotEnc**
  - **sentBytesEnc**, **sentBytesNotEnc**
  - **numConnsEnc**, **numConnsNotEnc**
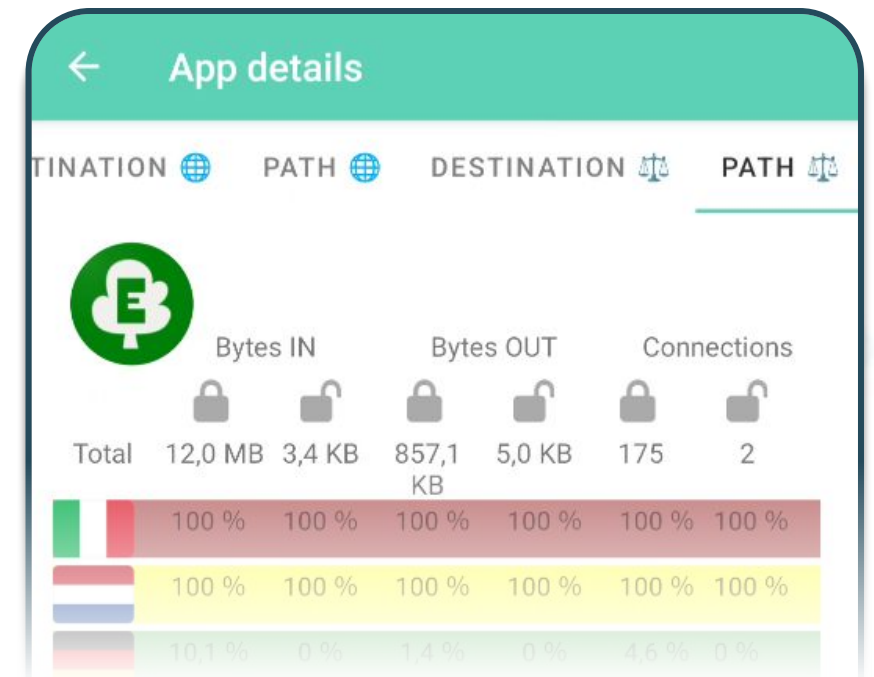
Manana Server

Collects the reports provided by the users

- Each report is received by an **ftp server** (running in its own virtual machine) and then moved to an archive outside the VM

- Periodically (30 min.) the **reports are aggregated** and newly updated statistics are generated using all the reports

- The newly generated stats are saved to a **SQLite database**

- The updated statistics are published by the **web server**

Manana Server: statistics

- For **each app** the computed statistics include:

  - For **each destination AS** and **destination country**
    - The total count of the entries in all the reports and the relative overall share
    - The share of data sent/received
    - The share of packets sent/received

  - For each **destination country**, a **safety rating** is shown (with same criteria adopted for the mobile app, based on V-Dem dataset)
  - Similarly, the entry count, share and safety rating of each **country traversed** according to the traceroute reports

NET MAKERS

Manana Server: statistics example- Gmail

Search App

| Gmail |

Statistics of: Gmail

Stats by Autonomous System Number

ASN Name	Entry Count	Entry Percentage	Bytes Sent Percentage	Bytes Rcvd Percentage	Packets Sent Percentage	Packets Rcvd Percentage
AS8075 - Microsoft Corporation	31	44.29%	39.60%	58.01%	45.86%	45.62%
AS396982 - Google LLC	34	48.57%	35.45%	33.60%	46.71%	47.24%
Unknown ASN	1	1.43%	0.33%	0.35%	0.21%	0.23%
AS15169 - Google LLC	4	5.71%	24.63%	8.05%	7.22%	6.91%

Stats by Destination Country

Country Name	Entry Count	Entry Percentage	Bytes Sent Percentage	Bytes Rcvd Percentage	Packets Sent Percentage	Packets Rcvd Percentage	Safety
Ireland	6	8.57%	2.16%	2.88%	4.46%	2.76%	DANGEROUS
Italy	19	27.14%	39.71%	36.64%	35.88%	37.10%	DANGEROUS
Netherlands	34	48.57%	35.45%	33.60%	46.71%	47.24%	NON-SECURE
Sweden	3	4.29%	7.20%	8.84%	3.18%	3.00%	DANGEROUS
France	7	10.00%	15.16%	17.69%	9.55%	9.68%	NON-SECURE
No Info	1	1.43%	0.33%	0.35%	0.21%	0.23%	UNKNOWN

Stats by Traversed Countries

Country Name	Entry Count	Percentage	Safety
Italy	1	25.00%	DANGEROUS
United States	1	25.00%	SECURE
Netherlands	2	50.00%	NON-SECURE

Manana Server: statistics example- Play Store

Search App

Google Play Store

Statistics of: Google Play Store

Stats by Autonomous System Number

ASN Name	Entry Count	Entry Percentage	Bytes Sent Percentage	Bytes Rcvd Percentage	Packets Sent Percentage	Packets Rcvd Percentage
AS15169 - Google LLC	29	96.67%	99.66%	99.86%	99.44%	99.42%
Unknown ASN	1	3.33%	0.34%	0.14%	0.56%	0.58%

Stats by Destination Country

Country Name	Entry Count	Entry Percentage	Bytes Sent Percentage	Bytes Rcvd Percentage	Packets Sent Percentage	Packets Rcvd Percentage	Safety
United Kingdom	19	63.33%	32.16%	8.29%	47.46%	40.94%	DANGEROUS
Italy	5	16.67%	60.99%	77.90%	42.94%	50.88%	DANGEROUS
United States	3	10.00%	4.47%	8.62%	5.65%	4.09%	SECURE
France	2	6.67%	2.04%	5.04%	3.39%	3.51%	NON-SECURE
No Info	1	3.33%	0.34%	0.14%	0.56%	0.58%	UNKNOWN

Stats by Traversed Countries

Country Name	Entry Count	Percentage	Safety
Italy	1	25.00%	DANGEROUS
United States	1	25.00%	SECURE
France	2	50.00%	NON-SECURE

Very early stage development

the prototype is demonstrating feasibility, not ready for common public yet (you can have a run for fun, stay tuned for updates)

- ONGOING DEVELOPMENT
  - code **cleansing**
  - **corner** cases (time-outs, missing info)
  - extensive **testing**
  - manage **issue** ticketing
  - process **feedback** about usability and privacy requirements
  - **documentation** in app, and on website

NET MAKERS

Improvement plan

Once the main functionalities have been finalized

- Planned
  - create domain add SSL certificates for MANANA server
  - publish on **F-Droid**
    (for independent vetting and auto-update of new versions)
  - analyze **crowdsourced data** for reporting
  - explore **different** criteria for country **information-safety classification**
  - analyze data for **anomalies** (geoIP / WHOIS artifacts)
    - design and implement **improvements of accuracy**

- Nice-to-have
  - provide a **direct comparison** personal-vs-crowdsourced statistics (to allow the user to compare her results with others / other not-installed apps)
  - allow the user to set her **custom country safety evaluation**

Manana Server running instance

Source code repository

MANANA 0.2.1-beta APK

http://143.225.229.154

https://codeberg.org/MANANA_project-UniNA

https://www.traffic.comics.unina.it/software/manana/manana-0.2.1-beta.apk

Contacts: **giuseppe.aceto@unina.it**
Follow the project on codeberg:
https://codeberg.org/MANANA_project-UniNA.rss

HAVE A TASTE!

WORK SHOP GARR 2024

NET MAKERS

ACKNOWLEDGEMENTS AND REFERENCES

Source code of the app and the server can be found here: https://codeberg.org/MANANA_project-UniNA

The **information-safety** classification criteria (ab)use the **dataset** published as
Coppedge, Michael, et al. 2024. "V-Dem Country-Year Dataset v14", Varieties of Democracy (V-Dem) Project.
https://doi.org/10.23696/mcwt-fr58

The **IP geolocation** dataset is "IP to Country Lite" Lite by DB-IP https://db-ip.com
licensed under a Creative Commons Attribution 4.0 International License

The **mobile app code** is based on PCAPdroid (traffic capture and analysis) by Emanuele Faranda
https://github.com/emanuele-f/PCAPdroid licensed under GPL3.0

The **traceroute code** is from the Mobiperf project by Google
https://github.com/Mobiperf licensed under Apache License 2.0

The **ftp server** is vsftpd https://security.appspot.com/vsftpd.html (GPL lincense)

The **statistics web server** is implemented in the Django framework
https://www.djangoproject.com (BSD license)

Backup slides

Path analysis via Traceroute

- Traceroute implementation by MobiPerf
  - Current traceroute implementation sends out <u>three ICMP probes</u> per TTL
  - One ping every 0.2s is the lower bound before some platforms require root to run ping.
  - reduced timeout w.r.t. MobiPerf default

Path-tracing limitations in MANANA

- reverse-path uncertainties
  - *Internet routing may be asymmetric*
  - *inbound traffic not guaranteed to follow the path of the probes*
  - *traceroute inherent limitation, we have to live with this*
- path stability assumptions
  - *the path towards each destination is sampled once per observation session*
  - *no big issue in our application scenario: no impact on AS- and country-paths inferred*
    - *Internet routes are expected to change at a lower frequency*
    - *ECMP is expected to be employed within an AS*
- ICMP vs TCP/UDP data probes
  - crafting probes of the same type of data packet would ensure to follow the very same path
  - in MANANA app, we are limited to ICMP to <u>avoid rooting</u> of the smartphone (not general audience)
- third-party IP addresses
  - the RFC1812 states that the source address of an ICMP error packet should correspond to the outgoing interface of the ICMP reply, rather than the interface on which the packet triggering the error was received
  - this behavior can cause a traceroute IP path to include addresses associated to interfaces not included in the path actually traversed
  - this may cause the inference of inaccurate AS paths
  - path-data cleaning (e.g., checking for AS paths shorter that 3 hops during post-processing?) can mitigate such issue
- AS-border identification
  - may impact length of the portion of the path within an AS, and thus proportions and stats
  - not a major issue in our case
  - https://dl.acm.org/doi/abs/10.1145/3278532.3278538
- IP Geofeeds
  - https://hal.science/hal-04663776/document
  - RFC 8805 (2021), draft 2013. info included in WHOIS data (inetnum objects)

traceroute: report.json example

```json
[
  {
    "AppName": "Root",
    "WhoisDestinationCoutries": [
      {
        "IpList": [
          {
            "8.8.8.8":
"2024-10-29T16:33:18.645+01:00"
          }
        ],
        "classificationValue": 3,
        "name": "US",
        "numConnsEnc": 0,
        "numConnsNotEnc": 2,
        "rcvdBytesEnc": 0,
        "rcvdBytesNotEnc": 196,
        "sentBytesEnc": 0,
        "sentBytesNotEnc": 135
      }
    ],
    "WhoisTraceCountries": {
      "8.8.8.8": {
        "numConnsEnc": 0,
        "numConnsNotEnc": 2,
        "orderedListCountry": [
          "NL",
          "US",
          "US"
        ],
        "rcvdBytesEnc": 0,
        "rcvdBytesNotEnc": 196,
        "result": [],
        "sentBytesEnc": 0,
        "sentBytesNotEnc": 135,
        "timeStamp":
"2024-10-29T16:38:20.207+01:00"
      }
    },
    "traceCountries": {
      "8.8.8.8": {
        "numConnsEnc": 0,
        "numConnsNotEnc": 2,
        "orderedListCountry": [
          "IT",
          "US",
          "US"
        ],
        "rcvdBytesEnc": 0,
        "rcvdBytesNotEnc": 196,
        "result": [
          {
            "hosts": [
              "192.168.0.1:"
            ],
            "rtt": 13.666666666666666
          },
          "sentBytesEnc": 0,
          "sentBytesNotEnc": 135,
          "timeStamp":
"2024-10-29T16:33:36.893+01:00"
```

Updating Data

The data updates from a new connection are:

- **AppStats,** this class is associated to an app, its values are updated depending on the value of the UID in the connection. If the data are unencrypted it will updated its **sentBytesNotEnc**, **rcvdBytesNotEnc numConnsNotEnc** otherwise its **sentBytesEnc rcvdBytesEnc**, **numConnsEnc.**

- **Destination Countries,** a list of countries in AppStats made by DstIP geolocation of every connection generated by the app. It updates data for every country in the same way.

- **traceCountries,** a list of IPs and its relative Traceroute data in AppStats. The data are updated before the Traceroute execution in the same way mentioned above.

- **countriesAD** and **traceCountriesAD** (respectively **WhoisDestinationCoutries** and **WhoisTraceCountries** from the previous slide), also present in Appstats and update the data in the same way.