## 采集器需求

### 域名信誉情报采集

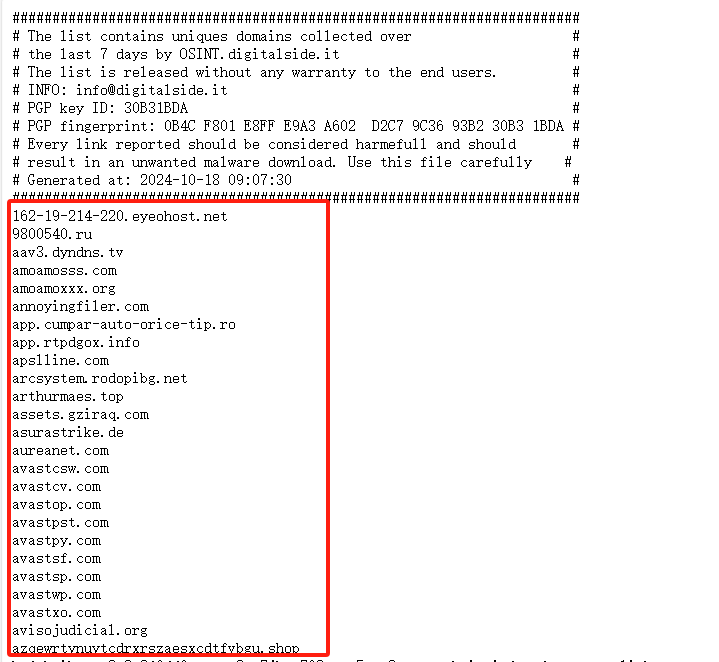
#### osint.digitalside.it（C2）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| domain | 1 | String | 第1列 | 域名 |
| threaten \_type | 1 | String | 8 | 攻击类型，可以属于多个类型 |
| credit\_level | 1 | Int | 5 | 信誉等级 |
| attack\_time | 1 | Date | 2024-10-18 09:07:30 | 恶意行为攻击时间 |
| threaten\_score | 1 | Int | 80 | 威胁指数，取值范围 0-100。 |
| Created\_time | 1 | Date | 情报入库时间 | 情报入库创建时间 |
| threaten\_source | 1 | Int | 80 | 情报来源分值 |
| source | 1 | string | 188 | 情报来源 |

###### 约束

1. 爬取时间：每小时定时爬取
2. topic：域名信誉
3. 地址：https://osint.digitalside.it/Threat-Intel/lists/latestdomains.txt
4. 固定取值部分：" threaten \_type ": "8"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":188
5. 注意：在页面拉取数据的时候，####框内的数据不用提取。取下图红框里的内容



###### Json示例

{

"domian": "aav3.dyndns.tv",

"threaten \_type": "8",

"credit\_level"：5,

"attack\_time":"2024-10-18 09:07:30",

"threaten\_score": 80,

"created\_time": "2025-01-14 00:00:00",

"threaten\_source":80,

"source":"188"

}

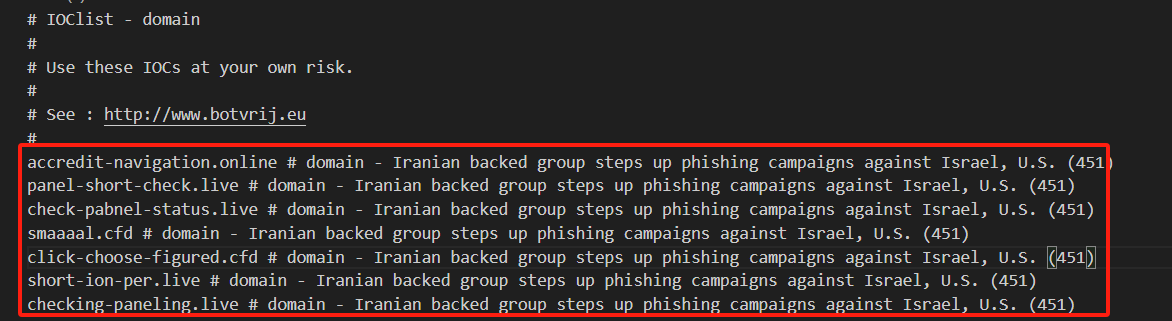
#### [www.botvrij.eu](http://www.botvrij.eu) (IOC)

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| domain | 1 | String | 第1列 | 域名 |
| threaten \_type | 1 | String | 42 | 攻击类型，可以属于多个类型 |
| credit\_level | 1 | Int | 5 | 信誉等级 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| threaten\_score | 1 | Int | 80 | 威胁指数，取值范围 0-100。 |
| Created\_time | 1 | Date | 情报入库时间 | 情报入库创建时间 |
| threaten\_source | 1 | Int | 80 | 情报来源分值 |
| source | 1 | string | 189 | 情报来源 |

###### 约束

1. 爬取时间：不定期爬取
2. topic：域名信誉
3. 地址：<https://www.botvrij.eu/data/ioclist.domain>
4. 固定取值部分：" threaten \_type ": "42"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":189
5. 注意：拉取文件前几行为页面信息，无需拉取，从有实际数据的地方开始。



###### 3）Json示例

{

"domian": "understandingthewar.org",

"threaten \_type": "42",

"credit\_level"：5,

"attack\_time":"2025-01-14 00:00:00",

"threaten\_score": 80,

"created\_time": "2025-01-14 14:38:18",

"threaten\_source":80,

"source":"189"

}

#### raw.githubusercontent.com（恶意ip）

###### 数据存储字段

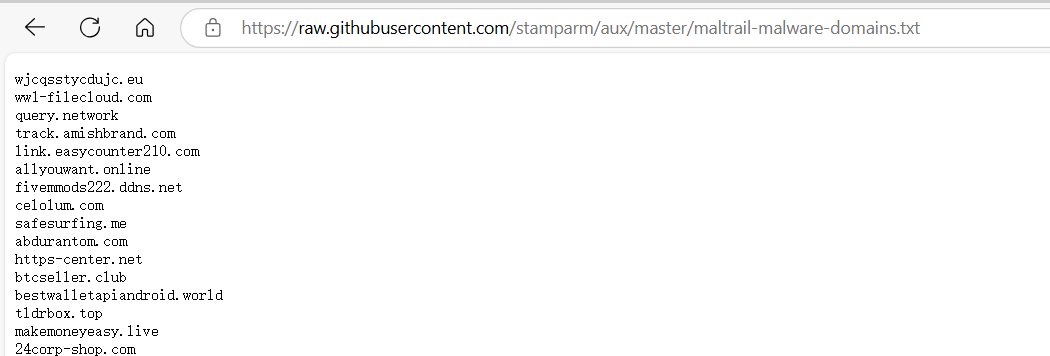
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| domain | 1 | String | 第1列 | 域名 |
| threaten \_type | 1 | String | 42 | 攻击类型，可以属于多个类型 |
| credit\_level | 1 | Int | 5 | 信誉等级 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| threaten\_score | 1 | Int | 80 | 威胁指数，取值范围 0-100。 |
| Created\_time | 1 | Date | 情报入库时间 | 情报入库创建时间 |
| threaten\_source | 1 | Int | 80 | 情报来源分值 |
| source | 1 | string | 13 | 情报来源 |

###### 约束

1. 爬取时间：每天爬取
2. topic：域名信誉
3. 地址：

http://raw.githubusercontent.com/stamparm/aux/master/maltrail-malware-domains.txt

1. 固定取值部分：" threaten \_type ": "42"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":" 13 "
2. 注意：页面无其他信息，只有域名信誉可以直接提取。



###### Json示例

{

"domian": "magentocore.net",

"threaten \_type": "42",

"credit\_level"：5,

"attack\_time":"2025-01-14 00:00:00",

"threaten\_score": 80,

"created\_time": "2025-01-14 14:48:26",

"threaten\_source":80,

"source ":"13"

}

#### torstatus.rueckgr.at（tor）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| domain | 1 | String | 第6列 | 域名 |
| threaten \_type | 1 | String | 44 | 攻击类型，可以属于多个类型 |
| credit\_level | 1 | Int | 5 | 信誉等级 |
| attack\_time | 1 | Date | 当天时间 | 恶意行为攻击时间 |
| threaten\_score | 1 | Int | 80 | 威胁指数，取值范围 0-100。 |
| Created\_time | 1 | Date | 情报入库时间 | 情报入库创建时间 |
| threaten\_source | 1 | Int | 80 | 情报来源分值 |
| source | 1 | string | 184 | 情报来源 |

###### 2）约束

1. 爬取时间：每天爬取
2. topic：域名信誉
3. 地址：https://torstatus.rueckgr.at/query\_export.php/Tor\_query\_EXPORT.csv
4. 固定取值部分：" threaten \_type ": "44"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":" 184"
5. 注意：打开链接会下拉一个excel，第四列有一个Uptime(days)字段，需要明确是否可以用于攻击时间。

###### 3）Json示例

{

"domian": "guybrush.jaehne.cloud",

"threaten \_type": "44",

"credit\_level"：5,

"attack\_time":"2025-01-14 00:00:00",

"threaten\_score": 80,

"created\_time": "2025-01-14 16:29:18",

"threaten\_source":80,

"source ":"84"

}

### IP信誉情报**采集**

#### vxvault.net（C2）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第4列 | ip地址 |
| attack\_time | 1 | Date | 第1列 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 8 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 59 | 情报来源 |

###### 2）约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：http://vxvault.net/ViriList.php
4. 固定取值部分：" threaten \_type ": "8"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":59
5. 注意：页面上攻击时间只显示月份与日。

###### 3）Json示例

{

"ip": "194.37.81.64",

"attack\_time":"01-02",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "8",

"threaten\_score ": 80,

"source ":"59"

}

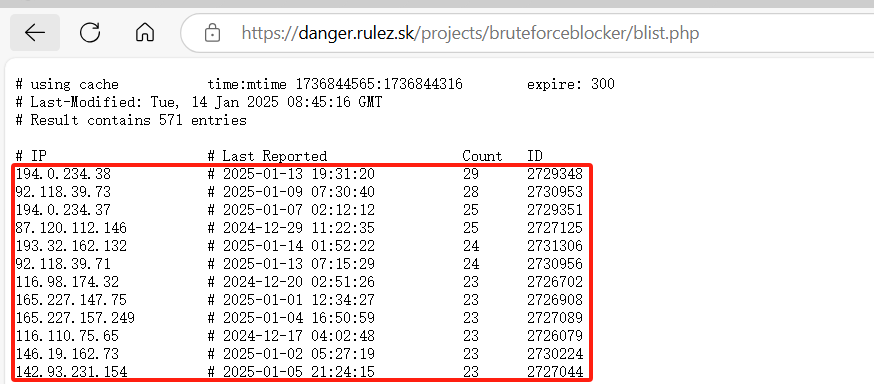
#### danger.rulez.sk

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 第2列 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 42 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 59 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：<http://danger.rulez.sk/projects/bruteforceblocker/blist.php>
4. 固定取值部分：" threaten \_type ": "42"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":59
5. 注意：页面前三行是说明信息，需从底6行开始获取有效信息



###### Json示例

{

"ip": "194.0.234.38",

"attack\_time":"2025-01-13 19:31:20",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "42",

"threaten\_score ": 80,

"source ":"59"

}

#### [www.dan.me.uk](http://www.dan.me.uk)

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 1 | 信誉等级 |
| threaten\_score | 1 | int | 30 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 44 | 恶意行为类型 |
| threaten \_source | 1 | Int |  | 恶意行为类型的威胁分数 |
| source | 1 | String |  | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：https://www.dan.me.uk/torlist/
4. 固定取值部分：" threaten \_type ": "44"，" credit\_level ": "1"，" threaten\_score ": "30"，" source ":
5. 注意：页面只有一列ip信息。

###### Json示例

{

"ip": "194.0.234.38",

"attack\_time":"2025-01-14 00:00:00",

"credit\_level "：1,

"threaten\_score ": 30，

"threaten \_type ": "44",

"threaten\_score ": 30,

"source ":""

}

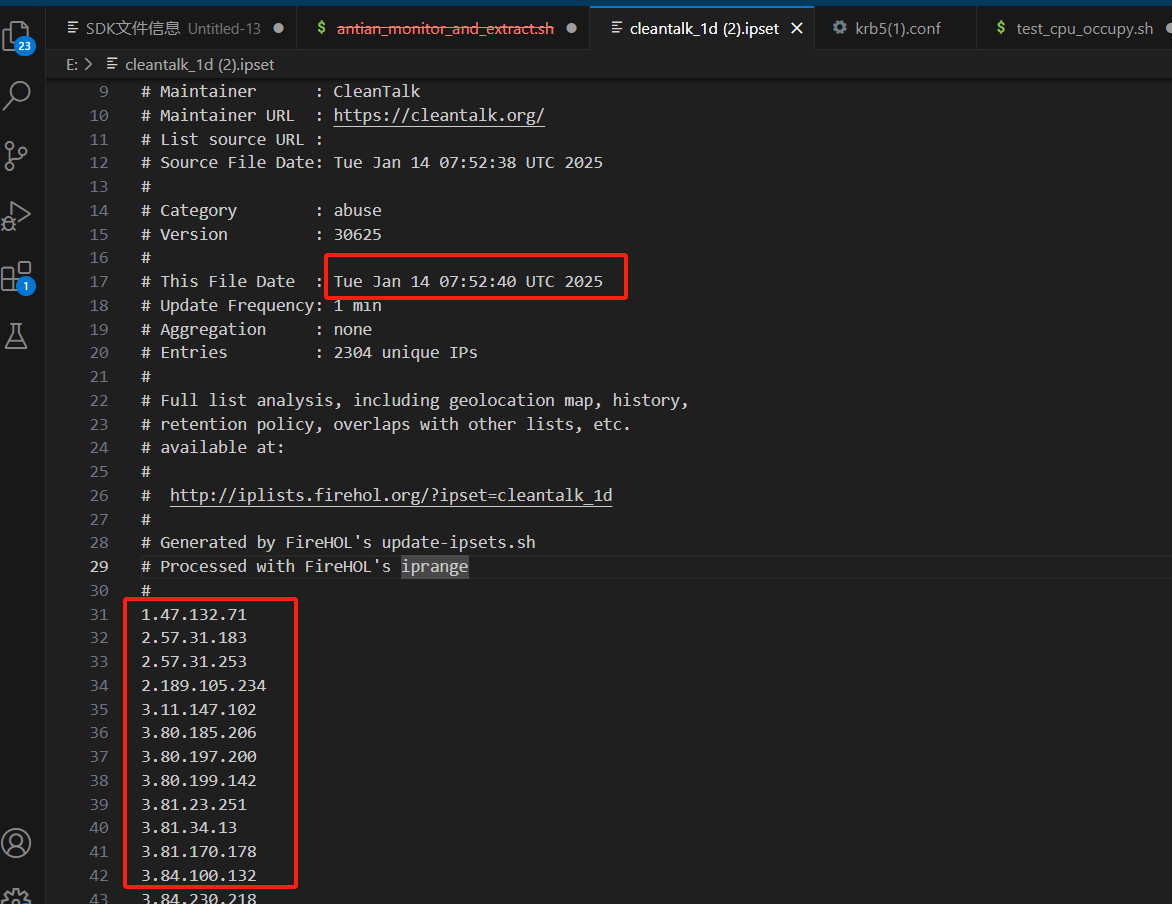
#### iplists.firehol.org

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 在上面的说明信息 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 40 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 185 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：https://iplists.firehol.org/files/cleantalk\_1d.ipset
4. 固定取值部分：" threaten \_type ": "40"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":185
5. 注意：无攻击时间，但文件上方说明信息中有文件生成时间，可以将这个当作攻击时间。IP信息在一段说明后。



###### Json示例

{

"ip": "3.80.185.206",

"attack\_time":"Tue Jan 14 07:52:40 UTC 2025",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "40",

"threaten\_score ": 80,

"source ":"185"

}

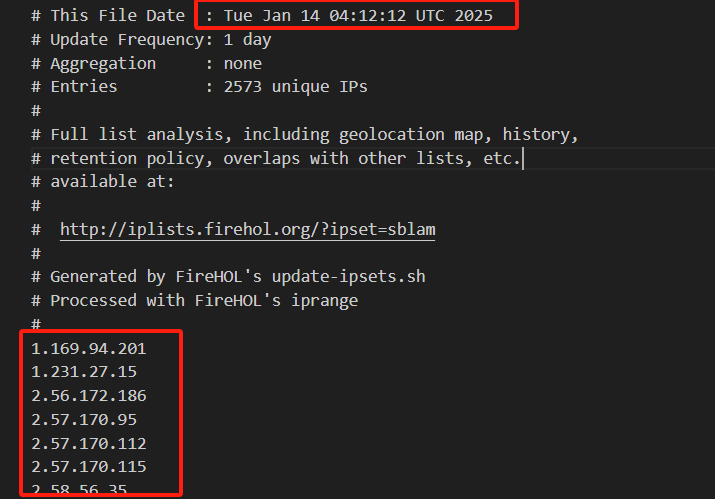
#### iplists.firehol.org

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 在上方的页面说明里 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 40 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 185 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：https://iplists.firehol.org/files/sblam.ipset
4. 固定取值部分：" threaten \_type ": "40"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":185
5. 注意：无攻击时间，但文件上方说明信息中有文件生成时间，可以将这个当作攻击时间。IP信息在一段说明后。



###### Json示例

{

"ip": "2.57.170.112",

"attack\_time":"Tue Jan 14 04:12:12 UTC 2025",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "40",

"threaten\_score ": 80,

"source ":"185"

}

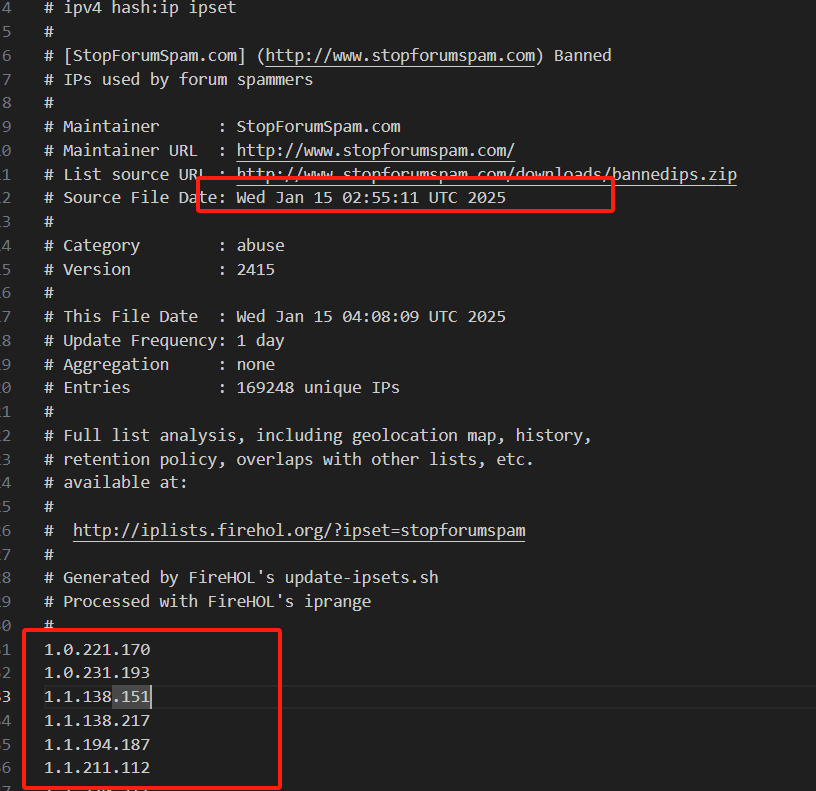
#### iplists.firehol.org

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 40 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 185 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：IP信誉
3. 地址：https://iplists.firehol.org/files/stopforumspam.ipset
4. 固定取值部分：" threaten \_type ": "40"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":185
5. 注意：在文件说明信息中有一个日期时间，可以看是否能使用。有效信息在说明信息后。



###### Json示例

{

"ip": "1.1.194.187",

"attack\_time":"2025-01-15 00:00:00",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "40",

"threaten\_score ": 80,

"source ":"185"

}

#### iplists.firehol.org （情报类型存疑 tor）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 46 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 185 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://iplists.firehol.org/files/tor\_exits.ipset
4. 固定取值部分：" threaten \_type ": "46"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":185
5. 注意：在文件说明信息中有一个日期时间，可以看是否能使用。有效信息在说明信息后。

###### Json示例

{

"ip": "5.255.98.151",

"attack\_time":"2025-01-21 00:00:00",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "46",

"threaten\_score ": 80,

"source ":"185"

}

#### stopforumspam.com

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 第3列 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 3 | 信誉等级 |
| threaten\_score | 1 | int | 60 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 0 | 恶意行为类型 |
| threaten \_source | 1 | Int | 60 | 恶意行为类型的威胁分数 |
| source | 1 | String | 54 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://stopforumspam.com/downloads/listed\_ip\_1\_all.zip
4. 固定取值部分：" threaten \_type ": "0"，" credit\_level ": "3"，" threaten\_score ": "60"，" source ":54
5. 注意：

###### Json示例

{

"ip": "1.1.223.77",

"attack\_time":"2025-01-20 06:50:21",

"credit\_level "：3,

"threaten\_score ": 60，

"threaten \_type ": "0",

"threaten\_score ": 60,

"source ":"54"

}

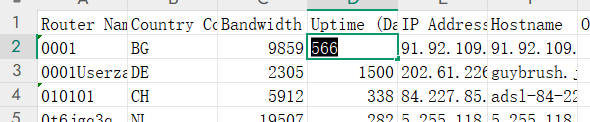
#### torstatus.rueckgr.at （情报类型存疑 tor）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第5列 | ip地址 |
| attack\_time | 1 | Date | 第4列 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 46 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 184 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://torstatus.rueckgr.at/query\_export.php/Tor\_query\_EXPORT.csv
4. 固定取值部分：" threaten \_type ": "46"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":184
5. 注意：文件中的时间是ip出现至今的天数，若不好操作就改成拉取当天0点



###### 3）Json示例

{

"ip": "91.92.109.43",

"attack\_time":"2025-01-21 00:00:00",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "46",

"threaten\_score ": 80,

"source ":"184"

}

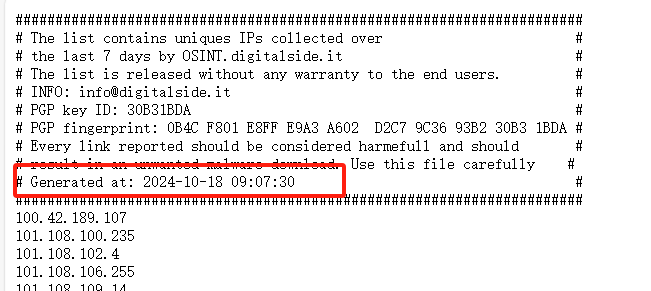
#### osint.digitalside.it

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 10 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 188 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://osint.digitalside.it/Threat-Intel/lists/latestips.txt
4. 固定取值部分：" threaten \_type ": "10"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":188
5. 注意：在顶端的说明信息中有一个文件生成时间，看看是否可以当作攻击时间使用。



###### Json示例

{

"ip": "100.42.189.107",

"attack\_time":""2024-10-18 09:07:30,

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "10",

"threaten\_score ": 80,

"source ":"188"

}

#### raw.githubusercontent.com （无法对应上情报源）

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当天0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 13 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String |  | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：

https://raw.githubusercontent.com/stamparm/ipsum/master/ipsum.txt

1. 固定取值部分：" threaten \_type ": "13"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":
2. 注意：

###### Json示例

{

"ip": "5.180.253.220",

"attack\_time":"2025-01-21 00:00:00",

"credit\_level "：5,

"threaten\_score ": 80，

"threaten \_type ": "13",

"threaten\_score ": 80,

"source ":""

}

#### myip.ms (威胁类型存疑)

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 第2列 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 1 | 信誉等级 |
| threaten\_score | 1 | int | 30 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 20 | 恶意行为类型 |
| threaten \_source | 1 | Int | 30 | 恶意行为类型的威胁分数 |
| source | 1 | String | 14 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://myip.ms/files/blacklist/general/latest\_blacklist.txt
4. 固定取值部分：" threaten \_type ": "20"，" credit\_level ": "1"，" threaten\_score ": "30"，" source ":14
5. 注意：

###### Json示例

{

"ip": "182.160.110.125",

"attack\_time":"2025-01-11",

"credit\_level "：1,

"threaten\_score ": 30，

"threaten \_type ": "20",

"threaten\_score ": 30,

"source ":"13"

}

#### reputation.alienvault.com

###### 数据存储字段

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数名称 | 级别 | 类型 | 对应字段 | 描述 |
| ip | 1 | String | 第1列 | ip地址 |
| attack\_time | 1 | Date | 当日0点 | 恶意行为攻击时间 |
| credit\_level | 1 | int | 5 | 信誉等级 |
| threaten\_score | 1 | int | 80 | 威胁指数，取值范围 0-100（根据该 IP 信誉、关联样本、历史记录评估得来）。 |
| threaten \_type | 1 | String | 2 | 恶意行为类型 |
| threaten \_source | 1 | Int | 80 | 恶意行为类型的威胁分数 |
| source | 1 | String | 59 | 情报来源 |

###### 约束

1. 爬取时间：每天定时爬取
2. topic：ipReputation
3. 地址：https://reputation.alienvault.com/reputation.data
4. 固定取值部分：" threaten \_type ": "2"，" credit\_level ": "5"，" threaten\_score ": "80"，" source ":59
5. 注意：

###### Json示例

{

"ip": "45.146.164.110",

"attack\_time":"2025-01-21 00:00:00",

"credit\_level "：3,

"threaten\_score ": 80，

"threaten \_type ": "2",

"threaten\_score ": 80,

"source ":"59"

}