Network Working Group
Internet-Draft
Intended status: Informational
Expires: October 2, 2019

A. Dulaunoy
A. Iklody
D. Servili
CIRCL

March 31, 2019

MISP galaxy format
draft-dulaunoy-misp-galaxy-format-06

Abstract

This document describes the MISP galaxy format which describes a simple JSON format to represent galaxies and clusters that can be attached to MISP events or attributes. A public directory of MISP galaxies is available and relies on the MISP galaxy format. MISP galaxies are used to add further informations on a MISP event. MISP galaxy is a public repository [MISP-G] [MISP-G-DOC] of known malware, threats actors and various other collections of data that can be used to mark, classify or label data in threat information sharing.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on October 2, 2019.

Table of Contents

1. Introduction

Sharing threat information became a fundamental requirements on the Internet, security and intelligence community at large. Threat information can include indicators of compromise, malicious file indicators, financial fraud indicators or even detailed information about a threat actor. Some of these informations, such as malware or threat actors are common to several security events. MISP galaxy is a public repository [MISP-G] of known malware, threats actors and various other collections of data that can be used to mark, classify or label data in threat information sharing.

In the MISP galaxy context, clusters help analysts to give more informations about their cybersecurity events, indicators or threats. MISP galaxies can be used for classification, filtering, triggering actions or visualisation depending on their use in threat intelligence platforms such as MISP [MISP-P].

1.1. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Format

A cluster is composed of a value (MUST), a description (OPTIONAL) and metadata (OPTIONAL).

Clusters are represented as a JSON [RFC4627] dictionary.

2.1. Overview

The MISP galaxy format uses the JSON [RFC4627] format. Each galaxy is represented as a JSON object with meta information including the following fields: name, uuid, description, version, type, authors, source, values, category.

name defines the name of the galaxy. The name is represented as a string and MUST be present. The uuid represents the Universally Unique IDentifier (UUID) [RFC4122] of the object reference. The uuid MUST be preserved. For any updates or transfer of the same object reference. UUID version 4 is RECOMMENDED when assigning it to a new object reference and MUST be present. The description is represented as a string and MUST be present. The uuid is represented as a string and MUST be present. The version is represented as a decimal and MUST be present. The type is represented as a string and MUST be present and MUST match the name of the galaxy file. The source is represented as a string and MUST be present. Authors are represented as an array containing one or more authors and MUST be present. The category is represented as a string and MUST be present and describes the overall category of the galaxy such as tool or actor.

Values are represented as an array containing one or more values and MUST be present. Values defines all values available in the galaxy.

2.2. values

The values array contains one or more JSON objects which represent all the possible values in the galaxy. The JSON object contains four fields: value, description, uuid and meta. The value is represented as a string and MUST be present. The description is represented as a string and SHOULD be present. The meta or metadata is represented as a JSON list and SHOULD be present. The uuid represents the Universally Unique IDentifier (UUID) [RFC4122] of the value reference. The uuid SHOULD can be present and MUST be preserved.

2.3. related

Related contains a list of JSON key value pairs which describe the related values in this galaxy cluster or to other galaxy clusters. The JSON object contains three fields, dest-uuid, type and tags. The

dest-uuid represents the target UUID which encompasses a relation of some type. The dest-uuid is represented as a string and MUST be present. The type is represented as a string and MUST be present and SHOULD be selected from the relationship types available in MISP objects [MISP-R]. The tags is a list of string which labels the related relationship such as the level of similarities, level of certainty, trust or confidence in the relationship, false-positive. A tag is represented in machine tag format which is a string an SHOULD be present.

```
"related": [ {
  "dest-uuid": "f873db71-3d53-41d5-b141-530675ade27a",
  "type": "similar",
  "tags": ["estimative-language:likelihood-probability=\"very-likely\""]
} ]
```

2.4. meta

Meta contains a list of custom defined JSON key value pairs. Users SHOULD reuse commonly used keys such as complexity, effectiveness, country, possible_issues, colour, motive, impact, refs, synonyms, status, date, encryption, extensions, ransomnotes, ransomnotes-filenames, ransomnotes-refs, suspected-victims, suspected-state-sponsor, type-of-incident, target-category, cfr-suspected-victims, cfr-suspected-state-sponsor, cfr-type-of-incident, cfr-target-category, attribution-confidence, payment-method, price wherever applicable.

refs, synonyms SHALL be used to give further informations. refs is represented as an array containing one or more strings and SHALL be present. synonyms is represented as an array containing one or more strings and SHALL be present.

date, status MAY be used to give time information about an cluster. date is represented as a string describing a time or period and SHALL be present. status is represented as a string describing the current status of the clusters. It MAY also describe a time or period and SHALL be present.

colour fields MAY be used at predicates or values level to set a specify colour that MAY be used by the implementation. The colour field is described as an RGB colour fill in hexadecimal representation.

complexity, effectiveness, impact, possible_issues MAY be used to give further information in preventive-measure galaxy. complexity is represented by an enumerated value from a fixed vocabulary and SHALL be present. effectiveness is represented by an enumerated value from

a fixed vocabulary and SHALL be present. impact is represented by an enumerated value from a fixed vocabulary and SHALL be present. possible_issues is represented as a string and SHOULD be present.

Example use of the complexity, effectiveness, impact, possible_issues fields in the preventive-measure galaxy:


```
{
  "meta": {
    "refs": [
      "http://www.windowsnetworking.com/kbase/WindowsTips/WindowsXP/AdminTips/
Customization/DisableWindowsScriptingHostWSH.html"
    ],
    "complexity": "Low",
    "effectiveness": "Medium",
    "impact": "Medium",
    "type": [
      "GPO"
    ],
    "possible_issues": "Administrative VBS scripts on Workstations"
  },
  "value": "Disable WSH",
  "description": "Disable Windows Script Host",
  "uuid": "e6df1619-f8b3-476c-b5cf-22b4c9e9dd7f"
}
```


country, motive MAY be used to give further information in threat-actor galaxy. country is represented as a string and SHOULD be present. motive is represented as a string and SHOULD be present.

Example use of the country, motive fields in the threat-actor galaxy:

```
{
  "meta": {
    "country": "CN",
    "synonyms": [
      "APT14",
      "APT 14",
      "QAZTeam",
      "ALUMINUM"
    ],
    "refs": [
      "http://www.crowdstrike.com/blog/whois-anchor-panda/"
    ],
    "motive": "Espionage",
    "attribution-confidence": 50
  },
  "value": "Anchor Panda",
  "description": "PLA Navy",
  "uuid": "c82c904f-b3b4-40a2-bf0d-008912953104"
}
```

encryption, extensions, ransomnotes, ransomnotes-filenames, ransomnotes-refs, payment-method, price MAY be used to give further information in ransomware galaxy. encryption is represented as a string and SHALL be present. extensions is represented as an array containing one or more strings and SHALL be present. ransomnotes is represented as an array containing one or more strings ans SHALL be present. ransomnotes-filenames is represented as an array containing one or more strings ans SHALL be present. ransomnotes-refs is represented as an array containing one or more strings ans SHALL be present.

Example use of the encryption, extensions, ransomnotes fields in the ransomware galaxy:

```
{
  "description": "Similar to Samas and BitPaymer, Ryuk is specifically used to
target enterprise environments. Code comparison between versions of Ryuk and
Hermes ransomware indicates that Ryuk was derived from the Hermes source code
and has been under steady development since its release. Hermes is commodity
ransomware that has been observed for sale on forums and used by multiple
threat actors. However, Ryuk is only used by GRIM SPIDER and, unlike Hermes,
Ryuk has only been used to target enterprise environments. Since Ryuk's
appearance in August, the threat actors operating it have netted over 705.80
BTC across 52 transactions for a total current value of $3,701,893.98 USD.",
  "meta": {
    "ransomnotes-filenames": [
      "RyukReadMe.txt"
    ],
    "ransomnotes-refs": [
      "https://www.crowdstrike.com/blog/wp-content/uploads/2019/01/RansomeNote-
fig3.png",
      "https://www.crowdstrike.com/blog/wp-content/uploads/2019/01/RansomeNote-
fig4.png"
    ],
    "refs": [
      "https://www.crowdstrike.com/blog/big-game-hunting-with-ryuk-another-
lucrative-targeted-ransomware/"
    ]
  },
  "uuid": "f9464c80-b776-4f37-8682-ffde0cf8f718",
  "value": "Ryuk ransomware"
}
```


source-uuid, target-uuid SHALL be used to describe relationships. source-uuid and target-uuid represent the Universally Unique IDentifier (UUID) [RFC4122] of the value reference. source-uuid and target-uuid MUST be preserved.

Example use of the source-uuid, target-uuid fields in the mitre-enterprise-attack-relationship galaxy:


```
      {
        "meta": {
          "source-uuid": "222fbd21-fc4f-4b7e-9f85-0e6e3a76c33f",
          "target-uuid": "2f1a9fd0-3b7c-4d77-a358-78db13adbe78"
        },
        "uuid": "cfc7da70-d7c5-4508-8f50-1c3107269633",
        "value": "menuPass (G0045) uses EvilGrab (S0152)"
      }
```

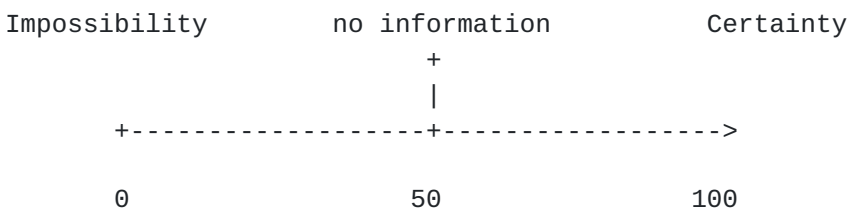

cfr-suspected-victims, cfr-suspected-state-sponsor, cfr-type-of-incident and cfr-target-category MAY be used to report information

gathered from CFR's (Council on Foreign Relations) [CFR] Cyber Operations Tracker. cfr-suspected-victims is represented as an array containing one or more strings and SHALL be present. cfr-suspected-state-sponsor is represented as a string and SHALL be present. cfr-type-of-incident is represented as a string or an array and SHALL be present. RECOMMENDED but not exhaustive list of possible values for cfr-type-of-incident includes "Espionage", "Denial of service", "Sabotage". cfr-target-category is represented as an array containing one or more strings ans SHALL be present. RECOMMENDED but not exhaustive list of possible values for cfr-target-category includes "Private sector", "Government", "Civil society", "Military".

Example use of the cfr-suspected-victims, cfr-suspected-state-sponsor, cfr-type-of-incident, cfr-target-category fields in the threat-actor galaxy:

```
{
  "meta": {
    "country": "CN",
    "refs": [
      "https://www.fireeye.com/blog/threat-research/2015/12/
the_eps_awakens.html",
      "https://www.cfr.org/interactive/cyber-operations/apt-16"
    ],
    "cfr-suspected-victims": [
      "Japan",
      "Taiwan"
    ],
    "cfr-suspected-state-sponsor": "China",
    "cfr-type-of-incident": "Espionage",
    "cfr-target-category": [
      "Private sector"
    ],
    "attribution-confidence": 50
  },
  "value": "APT 16",
  "uuid": "1f73e14f-b882-4032-a565-26dc653b0daf"
},
```

attribution-confidence MAY be used to indicate the confidence about an attribution given by country or cfr-suspected-state-sponsor. attribution-confidence is represented on a scale from 0 to 100, where 50 means "no information", the values under 50 mean "probably not, almost certainly not to impossibility", the values above 50 means "from probable, almost certain to certainty" and SHALL be present if country or cfr-suspected-state-sponsor are present.



3. JSON Schema

The JSON Schema [JSON-SCHEMA] below defines the overall MISP galaxy formats. The main format is the MISP galaxy format used for the clusters.

3.1. MISP galaxy format - galaxy

```
{
  "$schema": "http://json-schema.org/schema#",
  "title": "Validator for misp-galaxies - Galaxies",
  "id": "https://www.github.com/MISP/misp-galaxies/schema_galaxies.json",
  "type": "object",
  "additionalProperties": false,
  "properties": {
    "description": {
      "type": "string"
    },
    "type": {
      "type": "string"
    },
    "version": {
      "type": "integer"
    },
    "name": {
      "type": "string"
    },
    "icon": {
      "type": "string"
    },
    "uuid": {
      "type": "string"
    },
    "namespace": {
      "type": "string"
    },
    "kill_chain_order": {
      "type": "object"
    }
  },
  "required": [
    "description",
    "type",
    "version",
    "name",
    "uuid"
  ]
}
```

3.2. MISP galaxy format - clusters

```
{
  "$schema": "http://json-schema.org/schema#",
  "title": "Validator for misp-galaxies - Clusters",
```

```
  "id": "https://www.github.com/MISP/misp-galaxies/schema_clusters.json",
  "type": "object",
  "additionalProperties": false,
  "properties": {
    "description": {
      "type": "string"
    },
    "type": {
      "type": "string"
    },
    "version": {
      "type": "integer"
    },
    "name": {
      "type": "string"
    },
    "uuid": {
      "type": "string"
    },
    "source": {
      "type": "string"
    },
    "category": {
      "type": "string
    },
    "values": {
      "type": "array",
      "uniqueItems": true,
      "items": {
        "type": "object",
        "additionalProperties": false,
        "properties": {
          "description": {
            "type": "string"
          },
          "value": {
            "type": "string"
          },
          "uuid": {
            "type": "string"
          },
          "related": {
            "type": "array",
            "additionalProperties": false,
            "items": {
              "type": "object"
            },
            "properties": {
```

```
            "dest-uuid": {
              "type": "string"
            },
            "type": {
              "type": "string"
            },
            "tags": {
              "type": "array",
              "uniqueItems": true,
              "items": {
                "type": "string"
              }
            }
          }
        },
        "meta": {
          "type": "object",
          "additionalProperties": true,
          "properties": {
            "type": {
              "type": "array",
              "uniqueItems": true,
              "items": {
                "type": "string"
              }
            },
            "complexity": {
              "type": "string"
            },
            "effectiveness": {
              "type": "string"
            },
            "country": {
              "type": "string"
            },
            "possible_issues": {
              "type": "string"
            },
            "colour": {
              "type": "string"
            },
            "motive": {
              "type": "string"
            },
            "impact": {
              "type": "string"
            },
            "refs": {
```

```
                "type": "array",
                "uniqueItems": true,
                "items": {
                  "type": "string"
                }
              },
              "synonyms": {
                "type": "array",
                "uniqueItems": true,
                "items": {
                  "type": "string"
                }
              },
              "status": {
                "type": "string"
              },
              "date": {
                "type": "string"
              },
              "encryption": {
                "type": "string"
              },
              "extensions": {
                "type": "array",
                "uniqueItems": true,
                "items": {
                  "type": "string"
                }
              },
              "ransomnotes": {
                "type": "array",
                "uniqueItems": true,
                "items": {
                  "type": "string"
                }
              }
            }
          }
        },
        "required": [
          "value"
        ]
      }
    },
    "authors": {
      "type": "array",
      "uniqueItems": true,
      "items": {
```

```
        "type": "string"
      }
    }
  },
  "required": [
    "description",
    "type",
    "version",
    "name",
    "uuid",
    "values",
    "authors",
    "source",
    "category
  ]
}
```

4. Acknowledgements

The authors wish to thank all the MISP community who are supporting the creation of open standards in threat intelligence sharing.

5. References

5.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC4122] Leach, P., Mealling, M., and R. Salz, "A Universally Unique IDentifier (UUID) URN Namespace", RFC 4122, DOI 10.17487/RFC4122, July 2005, <https://www.rfc-editor.org/info/rfc4122>.

[RFC4627] Crockford, D., "The application/json Media Type for JavaScript Object Notation (JSON)", RFC 4627, DOI 10.17487/RFC4627, July 2006, <https://www.rfc-editor.org/info/rfc4627>.

5.2. Informative References

[CFR] CFR, "Cyber Operations Tracker - Council on Foreign Relations", 2018, <https://www.cfr.org/interactive/cyber-operations>.

[JSON-SCHEMA]
"JSON Schema: A Media Type for Describing JSON Documents", 2016, <https://tools.ietf.org/html/draft-wright-json-schema>.

[MISP-G] MISP, "MISP Galaxy - Public Repository", <https://github.com/MISP/misp-galaxy>.

[MISP-G-DOC]
MISP, "MISP Galaxy - Documentation of the Public Repository", <https://www.misp-project.org/galaxy.html>.

[MISP-P] MISP, "MISP Project - Malware Information Sharing Platform and Threat Sharing", <https://github.com/MISP>.

[MISP-R] MISP, "MISP Object Relationship Types - common vocabulary of relationships", <https://github.com/MISP/misp-objects/tree/master/relationships>.

Authors' Addresses

Alexandre Dulaunoy
Computer Incident Response Center Luxembourg
16, bd d'Avranches
Luxembourg L-1611
Luxembourg

Phone: +352 247 88444
Email: alexandre.dulaunoy@circl.lu

Andras Iklody
Computer Incident Response Center Luxembourg
16, bd d'Avranches
Luxembourg L-1611
Luxembourg

Phone: +352 247 88444
Email: andras.iklody@circl.lu

Deborah Servili
Computer Incident Response Center Luxembourg
16, bd d'Avranches
Luxembourg L-1611
Luxembourg

Phone: +352 247 88444
Email: deborah.servili@circl.lu