ISO/IEC 30184

Edition 1.0 2024-12

INTERNATIONAL STANDARD

colour inside

Internet of things (IoT) – Autonomous IoT object identification in a connected home – Requirements and framework

ISO/IEC 30184:2024-12(en)

IEC Secretariat
3, rue de Varembé
CH-1211 Geneva 20
Switzerland

Tel.: +41 22 919 02 11
info@iec.ch
www.iec.ch

ISO/IEC 30184

Edition 1.0 2024-12

INTERNATIONAL STANDARD

colour inside

Internet of things (IoT) – Autonomous IoT object identification in a connected home – Requirements and framework

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 35.020

ISBN 978-2-8327-0042-6

CONTENTS

INTERNET OF THINGS (IoT) – AUTONOMOUS IoT OBJECT IDENTIFICATION IN A CONNECTED HOME – REQUIREMENTS AND FRAMEWORK

FOREWORD

1) ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

2) The formal decisions or agreements of IEC and ISO on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC and ISO National bodies.

3) IEC and ISO documents have the form of recommendations for international use and are accepted by IEC and ISO National bodies in that sense. While all reasonable efforts are made to ensure that the technical content of IEC and ISO documents is accurate, IEC and ISO cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.

4) In order to promote international uniformity, IEC and ISO National bodies undertake to apply IEC and ISO documents transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC and ISO document and the corresponding national or regional publication shall be clearly indicated in the latter.

5) IEC and ISO do not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC and ISO marks of conformity. IEC and ISO are not responsible for any services carried out by independent certification bodies.

6) All users should ensure that they have the latest edition of this document.

7) No liability shall attach to IEC and ISO or their directors, employees, servants or agents including individual experts and members of its technical committees and IEC and ISO National bodies for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this ISO/IEC document or any other IEC and ISO documents.

8) Attention is drawn to the Normative references cited in this document. Use of the referenced publications is indispensable for the correct application of this document.

9) IEC and ISO draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). IEC and ISO take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, IEC and ISO had not received notice of (a) patent(s), which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at https://patents.iec.ch and www.iso.org/patents. IEC and ISO shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 30184 has been prepared by subcommittee 41: Internet of Things and Digital Twin, of ISO/IEC joint technical committee 1: Information technology.

The text of this International Standard is based on the following documents:

Draft	Report on voting
JTC1-SC41/453/FDIS	JTC1-SC41/469/RVD

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this International Standard is English.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1, and the ISO/IEC Directives, JTC 1 Supplement available at www.iec.ch/members_experts/refdocs and www.iso.org/directives.

IMPORTANT – The "colour inside" logo on the cover page of this document indicates that it contains colours which are considered to be useful for the correct understanding of its contents. Users should therefore print this document using a colour printer.

INTRODUCTION

The IoT environment has become widespread, dynamic, and complex, and is constantly evolving. IoT objects and their associations to users, or to other objects, should be identified. Current identification approaches rely on proper device categorization based on pre-determined taxonomies. Once categorized, devices advertise themselves to the network. When new types of IoT objects emerge, the taxonomy is renewed and new IDs are assigned.

As a complement to existing solutions, this document simplifies the requirements imposed on devices through the adoption of an autonomous procedure. This method reduces the need for detailed classification, standardization, and certification of device types by eliminating the need for devices to self-identify and advertise.

This document focuses on the requirements and the framework for autonomous identification of IoT objects, especially in connected home environments. The objects in this document include IoT devices and applications. The IoT object identification is to identify the IoT object type and the associations among the IoT objects.

Inspecting data patterns produced by IoT objects allows for autonomous type and association identification. The data patterns may be inspected if the IoT object has given explicit consent. The data patterns to be inspected can be a selected feature from the raw data such as the port number and protocol number. An accumulated feature set over time can also be used – minimum or maximum packet size, average input rate, average inter-arrival times of packets, and so on – if the IoT object gives explicit consent to allow the collection and storage of such data.

By doing so, the need for detailed classification, standardization, and certification of object types is reduced; and devices are relieved from the burdens of identifying and advertising themselves. It will motivate and spread the development of new types of IoT objects. Developments towards heterogeneous IoT objects will enable increased protections for devices and users against malicious attacks, hazards from malfunctions, or health-related critical issues.

INTERNET OF THINGS (IoT) – AUTONOMOUS IoT OBJECT IDENTIFICATION IN A CONNECTED HOME – REQUIREMENTS AND FRAMEWORK

1 Scope

This document specifies the following items for the autonomous IoT object identification in a connected home:

- requirements;
- architecture, functional entities and interfaces;
- operation procedures.

Information model formats, data formats, and identifier assignment are out of scope of this document.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

- IEC Electropedia: available at https://www.electropedia.org/
- ISO Online browsing platform: available at https://www.iso.org/obp

3.1
autonomous IoT object identification
identification of the *IoT object* (3.11) type and the associations among the IoT objects with limited human intervention

3.2
home
physical structure used as a dwelling place

EXAMPLE A house or an apartment.

Note 1 to entry: A home can be an individual building, part of a larger building or more than one building.

Note 2 to entry: A home can include small business premises, e.g. nursing homes and home offices.

[SOURCE: ISO/IEC 11801-4:2017 [1], 3.1.5, modified – Note 2 to entry has been added.]

3.3
connected home
home that is equipped with a home network

3.4
home network
internal network for information transport in a home or on business premises of similar complexity, providing defined access points and using one or more media in any topology

3.5
fingerprint
digital fingerprint
technology that deploys algorithms that analyse a large number of technical characteristics and settings on devices to generate unique identifiers that can identify a specific computing device producing a machine ID, and which can be personally identifiable

Note 1 to entry: In this document, a fingerprint is a selection of features. It can be an accumulated set of features over time.

[SOURCE: ISO 19731:2017 [2], 3.17, modified – Note 1 to entry has been replaced and Note 2 to entry has been deleted.]

3.6
feature
<machine learning> measurable property of an object or event with respect to a set of characteristics

Note 1 to entry: Features play a role in training and prediction.

Note 2 to entry: Features provide a machine-readable way to describe the relevant objects. As the algorithm will not go back to the objects or events themselves, feature representations are designed to contain all useful information.

[SOURCE: ISO/IEC 23053:2022 [3], 3.3.3]

3.7
ground truth
value of the target variable for a particular item of labelled input data

Note 1 to entry: The term ground truth does not imply that the labelled input data consistently corresponds to the real-world value of the target variables.

[SOURCE: ISO/IEC 22989:2022 [4], 3.2.7]

3.8
identifier
information that unambiguously distinguishes one entity from other entities in a given identity context

[SOURCE: ISO/IEC 20924:2024 [5], 3.1.19]

3.9
IoT application
software functional element specific to the solution of a problem in the IoT environment

Note 1 to entry: An application can be distributed among resources and can communicate with other applications.

[SOURCE: IEC 61800-7-1:2015 [6], 3.2.2, modified – The term and definition have been made specific to the IoT environment.]

3.10
IoT device

endpoint that interacts with the physical world through sensing or actuating

Note 1 to entry: An IoT device can be a sensor or an actuator.

[SOURCE: ISO/IEC 20924:2024 [5], 3.2.11]

3.11
IoT object
IoT device (3.10) and *IoT application* (3.9)

3.12
IoT system
system providing functionalities of IoT

Note 1 to entry: An IoT system can include, but not be limited to, *IoT devices* (3.10), IoT gateways, sensors, and actuators.

[SOURCE: ISO/IEC 20924:2024 [5], 3.2.15]

3.13
machine learning
ML
process of optimizing *model parameters* (3.15) through computational techniques, such that the *model*'s (3.14) behaviour reflects the data or experience

[SOURCE: ISO/IEC 22989:2022 [4], 3.3.5]

3.14
model
physical, mathematical or otherwise logical representation of a system, entity, phenomenon, process or data

[SOURCE: ISO/IEC 22989:2022 [4], 3.1.23]

3.15
model parameter
parameter
internal variable of a *model* (3.14) that affects how it computes its outputs

Note 1 to entry: Examples of parameters include the weights in a neural network and the transition probabilities in a Markov model.

[SOURCE: ISO/IEC 22989:2022 [4], 3.3.8]

3.16
meta-data
data that define and describe other data

[SOURCE: ISO/TR 3985:2021 [7], 3.10]

3.17
personally identifiable information
PII
information that can be used in a given context to identify, contact, or locate a single person, or to identify an individual in context

[SOURCE: ISO 19414:2020 [8], 3.1]

3.18
profile
set of attributes generated from one or more *fingerprints* (3.5) that represents characteristics of an IoT object

3.19
test data
evaluation data
data used to assess the performance of a final *model* (3.14)

Note 1 to entry: Test data are disjoint from *training data* (3.20) and *validation data* (3.21).

[SOURCE: ISO/IEC 22989:2022 [4], 3.2.14]

3.20
training data
data used to train a machine learning model

[SOURCE: ISO/IEC 22989:2022 [4], 3.3.16]

3.21
validation data
development data
data used to compare the performance of different candidate models

Note 1 to entry: Validation data are disjoint from *test data* (3.19) and generally also from *training data* (3.20). However, in cases where there are insufficient data for a three-way training, validation and test set split, the data are divided into only two sets – a test set and a training or validation set. Cross-validation or bootstrapping are common methods for then generating separate training and validation sets from the training or validation set.

Note 2 to entry: Validation data can be used to tune hyperparameters or to validate some algorithmic choices, up to the effect of including a given rule in an expert system.

[SOURCE: ISO/IEC 22989:2022 [4], 3.2.15]

4 Abbreviated terms

CF	central functions
CHCE	connected home control entity
COIF	central IoT object identification function
FEF	feature extraction function
FPGF	fingerprint and profile generation function
ID	identifier
LF	local functions
LOIF	local IoT object identification function
LPM	local policy manager
ODF	IoT object discovery function
PAF	policy application function
PDM	policy database manager
PII	personally identifiable information

5 Overview

As a complement to existing IoT unique device identification solutions, which remain crucial, like in cyber-physical forensics investigations, this document defines requirements, architecture, functional entities, and operation procedures for the autonomous identification and discovery of IoT devices and applications on the IoT devices. Inspecting data and patterns that IoT objects (3.11) produce is required in order to autonomously identify types and association of IoT objects.

This document focuses on the requirements and the framework for autonomous identification of IoT objects, especially in connected home environments. The IoT object identification is to identify the IoT object type and the associations among the IoT objects.

There can be architectures or operation procedures for similar purposes, which are different from those specified in this document.

The structure of this document is as follows.

Clause 6 specifies the requirements for the autonomous IoT object identification. The requirements are divided into major system capabilities and requirements.

Clause 7 specifies the architecture, the functional entities, and the interfaces between the entities. The architecture includes the local functions located at connected home servers and the central functions that are placed in cloud service platforms.

Clause 8 specifies the operation mechanisms and procedures for the autonomous IoT object identification, which involves participation and communication among end devices, connected home servers, and a cloud service platform.

6 Requirements

6.1 General description

This Clause 6 specifies the requirements for the autonomous IoT object identification. The requirements are divided into major system capabilities and system requirements.

6.2 Major system capabilities

IoT systems with the IoT object identification function shall have the following capabilities.

a) Autonomy: IoT systems with the IoT object identification function shall be able to operate with limited human intervention. The framework for this capability is covered in 7.2.

b) Scalability: IoT systems with the IoT object identification function shall be able to manage a large number of device and application types and learn to identify new types as they emerge. The framework for this capability is covered in 7.2.

c) Stability: IoT systems with the IoT object identification function shall be able to work consistently and effectively regardless of the lifecycle stage (e.g. whether in induction or normal operation stage) or operational mode (e.g. whether in standby or active mode) of the target IoT object or application. The framework for this capability is covered in 7.2.

d) Privacy: IoT systems with the IoT object identification function shall protect PII. The framework for this capability is covered in 7.2.2, 8.4, and 8.5.

6.3 System requirements and recommendations

This subclause 6.3 specifies the requirements and recommendations for systems with IoT object identification function. The requirements and recommendations are categorized based on the major system capabilities defined in 6.2.

a) Requirements and recommendations for autonomy of IoT object identification function.

   1) It shall be able to identify the type of IoT object. The type of an IoT object is an indication for the object, which can be determined based on the purpose, manufacturer, model number, version of firmware or software, or any combination of these attributes.

   2) It should be able to inspect the transferred data from unidentified IoT objects. Data including packet headers of all the layers, and unencrypted packet payload should be monitored.

   3) It should be able to extract selected features from the inspected data. Features from multiple packets over time should be able to be extracted.

   4) It shall be able to summarize accumulated features into a brief format, which is suitable as an input to the identification procedures.

b) Requirements and recommendations for scalability of IoT object identification function.

   1) It should be able to discover IoT objects and their identity through identity information exchange between the IoT objects and the functional entity responsible for the identification. Identity information and features of an IoT object obtained by this procedure should be used for accurate identification of other related IoT objects.

   2) It should be able to collect the essence of features of IoT objects from multiple connected homes and send them to a central computing system, such as a cloud or fog platform.

   3) It shall be able to identify IoT objects on a central computing system.

c) Requirements for stability of IoT object identification function.

   1) It shall be able to collect data transferred by IoT objects and identify data contents generated by IoT objects, if the consent is granted for the collection and transference of data from IoT objects.

   2) It shall provide a secure and safe communication environment for transmitting the features or fingerprints in the procedure of identifying IoT objects.

d) Requirements and recommendations for privacy in IoT object identification function.

   1) It shall give IoT objects the option to allow for the collection and processing of IoT object data.

   2) It should be able to identify the IoT object's allowance level to the identification procedure. IoT objects can be categorized into

      i) identification allowed, or

      ii) identification allowed but without communications to the central platform outside of the connected home, or

      iii) identification not allowed.

      The allowance level should be identified with explicit advertisements from IoT objects. Once the allowance level is identified, the identification procedure on the IoT object should be restricted, if necessary. If the allowance level cannot be identified, then the IoT system should consider the object to be of category iii).

   3) The IoT objects should be able to define the frequency and size of data transfer to the central platform for autonomous identification.

   4) It shall be able to protect PII, personal status information, and personal behavioural patterns, which can be collected, stored, and exchanged with the central processing entity usually located in the cloud platform.

NOTE Regional regulatory requirements can exist that define security and privacy.

7 Architecture

7.1 General description

This Clause 7 defines architecture, functional entities, and interfaces between the functional entities necessary to satisfy the requirements specified in Clause 6. The functional entities are divided into the local functions and the central functions. The former are placed within the connected home where devices and users are located. The latter are placed within a cloud service platform. The IoT object identification functions consider the security and privacy issues, but these are not the major objectives. It is assumed that a connected home control entity provides necessary control functions to the connected home, while communicating with the local functions of the IoT object identification.

7.2 Functional entities

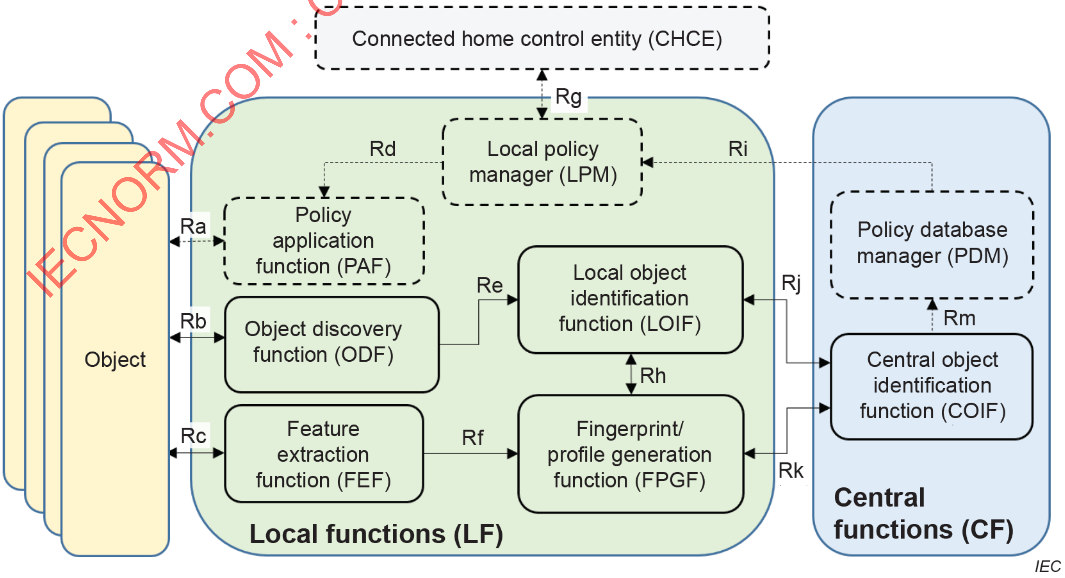
7.2.1 General

This subclause 7.2 defines functional entities that shall provide necessary functions to realize the autonomous IoT object identification. In 7.2, the term object is interchangeable with the term IoT object.

Figure 1 depicts a typical architecture with the local functions (LF) in a connected home and the central functions (CF) located usually in a cloud platform. There can be different architectures for similar purposes. The CF includes the central IoT object identification function (COIF) and the policy database manager (PDM). The LF includes the fingerprint and profile generation function (FPGF), the feature extraction function (FEF), the IoT object discovery function (ODF), the local IoT object identification function (LOIF), the local policy manager (LPM), and the policy application function (PAF). The connected home control entity (CHCE) represents the main control and management functions of a connected home.

The PAF, LPM, PDM, and CHCE are out of scope of this document. The PAF, LPM, PDM, CHCE, and the corresponding reference points in Figure 1 are informative.

The LF can be located within a gateway router of a connected home.



Connected home control entity (CHCE)
Rg
Rd
Local policy manager (LPM)
Ri
Policy application function (PAF)
Ra
Policy database manager (PDM)
Re
Local object identification function (LOIF)
Rj
Rb
Object discovery function (ODF)
Rm
Object
Rh
Central object identification function (COIF)
Rc
Feature extraction function (FEF)
Rf
Fingerprint/ profile generation function (FPGF)
Rk
Local functions (LF)
Central functions (CF)
IEC

NOTE The dotted entities and reference points are out of scope of this document. Functional entities are specified in 7.2.2 to 7.2.9.

Figure 1 – A typical architecture for autonomous IoT object identification

7.2.2 Feature extraction function

The FEF inspects the packets from the objects. The features can include data obtained from the physical layer header up to the application layer protocol headers. The accumulated features from the packets – packet inter-arrival times, packet length distributions, protocol-specific handshaking procedures, etc. – can be later processed to become the fingerprints.

Features are not usually involved directly with the user-level application payload data. However, when extracting such user data from IoT objects is necessary, for example images from surveillance cameras, PII shall not be extracted or collected. There can also be situations in which a natural person is identifiable even if there is no single attribute that uniquely identifies them. This is the case where a combination of several attributes taken together distinguishes this natural person from other natural persons. Whether or not a natural person is identifiable on the basis of a combination of attributes can also be dependent on the specific domain. For instance, the combination of the attributes gender, age, and occupation can be sufficient to identify a natural person within a particular home; see ISO/IEC 29100:2024 [9][1]. Therefore, access to the PII or any related information shall be controlled by actively disguising such combinations, or by not collecting some of these crucial personal attributes. The collection of PII data or attributes of a natural person shall not be used in the IoT object identification procedure. These data shall not be accessed and shall not be used.

7.2.3 Fingerprint and profile generation function

The FPGF generates the fingerprint and the profile of an IoT object based on the features extracted from the FEF. Fingerprints for identification can be changed over time. An optimization process that runs in COIF can decide different fingerprints and profiles as time passes. The FPGF shall be able to alter generation targets accordingly. This information shall be autonomously generated with minimum human intervention, in accordance with major system capability 6.2 a) Autonomy.

7.2.4 IoT object discovery function

The ODF exchanges the identity information of IoT objects, based on the traditional query and response or advertising procedures, e.g. as defined in ISO/IEC 30118-1:2021 [10]. The obtained information on the object types and resource types, combined with the matching fingerprint information, can be used as the ground truth or labelled data in machine learning-based IoT object identification procedures both in the LOIF and the COIF. However, the ODF is optional, and the IoT object identification functions should achieve their goals, irrespective of the existence of the ODF.

7.2.5 Local IoT object identification function

The LOIF identifies the IoT object, based on the information provided by the FPGF and the ODF, such as the matching information between the FPGF and the IoT object identity. The LOIF can use various identification mechanisms, including machine learning. It is also provided by the COIF regarding machine learning model deployment. These two identification functional entities can co-work with the knowledge distillation technique. The LOIF can customize the model deployed from the COIF, in accordance with local specifics. In addition, due to privacy issues and transmission cost, manually annotated data for training in the machine learning models are usually gradually collected and archived in separate sites. The LOIF is optional. However, the IoT object identification functions should achieve their goals, irrespective of the existence of the LOIF.

[1] Numbers in square brackets refer to the Bibliography.

7.2.6 Local policy manager

The LPM obtains the policy applicable to the IoT objects and the connected home from the PDM. The LPM distributes this information to the PAF and the CHCE.

LPM is outside the scope of this document and is only included for informative purposes.

7.2.7 Policy application function

The PAF enforces the policies that are related to the identification of the IoT object, through the CHCE. The CHCE represents the main control and management functions of a connected home and decides the policy to be enforced. The identification related policies are managed and suggested by the PDM.

PAF is outside the scope of this document and is only included for informative purposes.

7.2.8 Central IoT object identification function

The COIF gathers the information on the IoT objects from FPGFs and LOIFs of multiple different LFs. The COIF, which is likely within a cloud platform, and the LOIF shall work together in a consistent and stable manner, in accordance with major system capability 6.2 c) Stability. The LOIF shall also utilize the COIF properly to distribute excessive load, generated from the connected home, in accordance with Major system capability 6.2 b) Scalability.

The following statements regarding the COIF operations are informative: The COIF can use various mechanisms including fingerprint matching, context analysis, and machine learning. The ground truth information from the ODFs can be used for supervised learning. The learned or the refined matching information on the fingerprint and profile can be fed back to the LOIFs. The central IoT object identification can be a set of multiple instances distributed over clouds and multiple "fogs". Federated learning among these distributed heterogeneous models is also possible. In this case the communication and work distribution among these instances are important.

7.2.9 Policy database manager

The PDM maintains the policies that can be applied to the IoT objects, for example policies to further identify the associations of an IoT object, once it has been identified with known vulnerabilities. Based on this information, the CHCE can configure the corresponding firewall rules in accordance with the security requirements of the IoT object or notify the intrusion detection system to isolate vulnerable IoT objects. Another example is the network resource allocation policies for IoT objects with high priority traffic. The PDM can run a cache within the LF, for example within the LPM. Note that the policies suggested by the PDM are not the final policies to be enforced by the CHCE. The CHCE may have its own policy rules that are decided by the other functional entities than IoT object identification functions.

PDM is outside the scope of this document and is only included for informative purposes. An exemplary policy enforcement operation, which involves LPM, PAM, and PDM, is described in Annex A.

7.3 Reference points

This subclause 7.3 defines the reference points between the functional entities specified in 7.2. A reference point refers to an interface, or equivalently a conceptual point of interaction, between functional entities. The locations of the reference points are depicted in Figure 1. The informative reference points are depicted with dotted lines.

The reference point Ra allows the PAF to be able to enforce the IoT objects in accordance with the policies decided for the IoT objects based on their types and the associations. Ra also allows the IoT objects to inform its enforcement states to the PAF. Ra is outside the scope of this document and is only included for informative purposes.

The reference point Rb allows the identity information exchange between the IoT objects and the ODF. Rb should be compatible with the existing protocols.

The reference point Rc allows the information gathering of the FEF through various techniques such as traffic monitoring and packet inspection.

The reference point Rd allows the policy distribution from the LPM to the PAF. Rd is outside the scope of this document and is only included for informative purposes.

The reference point Re allows the ODF to provide the IoT object identity information to the LOIF.

The reference point Rf allows the FEF to provide a selection of the extracted features to the FPGF. Rf also allows the selection information to be provided by the FPGF to the FEF.

The reference point Rg allows the policy suggestion from the LPM to the CHCE. It also allows the identification related policy to be enforced to the IoT objects through the PAF. Rg is outside the scope of this document and is only included for informative purposes.

The reference point Rh allows the FPGF to provide the requested fingerprint or profile information to the LOIF. Rh also allows the fingerprint or profile generation rule to be sent from the LOIF to the FPGF.

The reference point Ri allows the PDM to provide the policies associated with IoT object identity to the LPM. Ri is outside the scope of this document and is only included for informative purposes.

The reference point Rj allows the multiple LOIFs to provide the requested identity information to the COIF. This information can include the labelled data to train and evaluate the machine learning models in the COIF. Rj also allows the information exchange between multiple LOIFs and the COIF regarding the trained machine learning model. Rj also allows the LOIFs and the COIF to cooperate for federated learning.

The reference point Rk can allow the multiple FPGFs to provide the requested fingerprint or profile information to the COIF for identification, based on the identification allowance level of the object. This information can include the data to train and evaluate the machine learning models in the COIF. Rk also allows the fingerprint and profile optimization information to be exchanged between the FPGFs and the COIF.

The reference point Rm allows the COIF to provide the identified IoT object information to the PDM. Rm is outside the scope of this document and is only included for informative purposes.

8 Operation procedure

8.1 Identifier

Device, or instance of application on a connected home, shall be identified by an identifier. An identifier itself does not identify an IoT object's type or association. As such, having an identifier does not mean that it is identified, according to the definition of identification in this document. An example of such an identifier could be a MAC address, an IP address, port number, or a combination of such information. A well-known identifier suitable for an IoT object, or for an IoT object instance, is the "flow ID". A flow is a set of packets of the same application, between a source and a destination, within a limited time period. It is usually characterized by the 5-tuplet (source IP, destination IP, source port, destination port, and IP protocol number). Although a device or an application can generate multiple flows at the same time, it still can be a simple and efficient method to indicate an IoT object.

8.2 Feature

Once the existence of a device or an application instance is revealed with an identifier, then the data from the target IoT object, or from other IoT objects that inspect the target, can be mapped with the identifier. Inspected raw meta-data are extracted into a feature set. An accumulated feature set over time can be further processed into a different form, which is defined in this document as a fingerprint.

Features can include unique hardware-specific characteristics like clock skew to identify unique network interface cards to identify individual IoT object instances. It can also include a large amount of packet headers from network traffic inspected over a long period of time for generating a proper fingerprint.

Another example of features is a set of network features extracted from the communication pattern of each flow. The features can be statistical data on these flows or flow type data: the average packet length, the minimum packet length, and the maximum length; the average of inter-arrival times of packets in a flow; the flow's size (number of packets in a flow); the protocols used by the flow (HTTP, HTTPS, SSDP, mDNS, TFTP, DHCP, DNS, NTP, BOOTP, TCP, or UDP).

Another example of features is a set of textual features extracted from IoT object's description shared in network payload in non-encrypted application layers: manufacturer name from MAC address; device name from DHCP information; manufacturer name, model, friendly name and type from XML description shared within UPnP messages during the discovery process; device local name, services names and types offered by the device from mDNS records; device OS, model and in some cases type from the user agent of the HTTP header, mainly used in mobile devices.

However, it is possible that this type of authentic features does not work correctly on malicious or abnormal IoT objects in the network. As such, a complementary solution for feature selection is necessary, for example based on unique behaviour. If an instant identification is not possible, malicious or abnormal IoT objects can be identified recursively using the identification procedure specified in this document.

8.3 Fingerprint and profile

Fingerprint and profile are information processed from the features monitored from IoT objects. An IoT object shall provide the user with an opt-in and opt-out mechanism for the collection and use of fingerprints.

A fingerprint is a careful selection of features. A fingerprint can also include data that are accumulated over time. For example, a fingerprint can include several selected header fields extracted from the first dozens of packets inspected from the object. As another example, a fingerprint can include textual information for IoT object identification, such as information modelled using a bag-of-words representation. A fingerprint can be even an inspected and processed image of a device, which is obtained through a different device such as a surveillance camera.

A fingerprint itself can be used as an input for IoT object identification functions, either for the LOIF or the COIF.

A fingerprint should be concise, because of its possible communication or processing overhead. For example, an image of a device can be too large to be transported to the CF in a cloud. A boundary object function (BOF) descriptor vector, which is a formalism that characterizes an object by extracting the boundary of an image, can be an example of a fingerprint as well.

A profile is a set of feature data associated with the fingerprint. It is used mainly for association identification. An example of profiles is context-based profile extracted from the events inferred by various sensing devices: the interval between the events, the types of sensed event, or the pattern of the events. These profiles can be used for identifying the association among IoT objects.

Deciding how to generate the fingerprint and profile manually can be a tedious and time-consuming process and can affect the real-time performance of the identification procedure. Sometimes the decision requires professional domain prior knowledge. Fingerprint and profile decision procedure itself can also be fully automated, for example using machine learning, with pre-processing techniques such as the traffic vectorization and the data augmentation.

8.4 IoT object type identification

Object type identification can be based on monitoring the communication behaviour of objects during the setup process. Object specific fingerprints are generated, which are mapped to object-types with the help of a machine learning-based classification model. Once an object type is identified, its associations are assessed further by the IoT identification functions.

A typical operation procedure for IoT object type identification is depicted in Figure 2.

a) Features are extracted from raw data transmitted by objects.

b) Features are sent to FGPF.

c) The COIF decides how the fingerprints and profiles are generated in the FPGF.

d) In the meantime, the ODF obtains by message exchanges with objects the ground-truth IoT object identities. This information is also sent to the LOIF. The LOIF can send this information to the COIF as well. This procedure is optional and in accordance with the allowance level.

e) Fingerprints and profiles are generated.

f) The generated fingerprints can be sent to the LOIF, COIF, or both, in accordance with the IoT object's allowance level.

g) Based on the identification method that the IoT system has selected, the COIF and LOIF can independently identify, or can adopt a cooperation model, such as federated or distillation-based learning.

h) In any case, the LOIF obtains identification results from the COIF and based on this result, finalizes the identification procedure.

The data exchanged between the local functions and central functions should be encrypted, in accordance with major system capability 6.2 d) Privacy.