ETSI TS 103 458 V1.1.1 (2018-06)

CYBER;
Application of Attribute Based Encryption (ABE) for PII and personal data protection on IoT devices, WLAN, cloud and mobile services - High level requirements

Reference
DTS/CYBER-0020

Keywords
access control, confidentiality, portability, privacy

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Cyber Security (CYBER).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document specifies high level requirements for the application of Attribute Based Encryption (ABE) to protect PII and personal data on IoT devices/services, cloud services, Wireless Local Area Networks and mobile services, where access to data has to be given to multiple parties and under different conditions. With a main focus on the confidentiality of data, including personal data and Personally Identifiable Information, the present document may help in supporting the General Data Protection Regulation [i.19].

The following use cases are described:

1) The Mobile use case describes a situation of user access from less trusted networks. The objective is to provide user identity protection preserving disclosure to unauthorized entity.

2) The federated WLAN use case where users can access different WLAN networks using their credentials - issued by different authorities/domains - while preserving their privacy.

3) Many Internet of Things use cases or edge scenarios where data access mechanisms are actioned either in the network or on the device.

4) The Cloud use case where a third party accesses personal data from the Cloud Service Provider.

The present document also provides recommendations on the ABE scheme to use for each use case.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

[1] ISO/IEC 17789:2014: "Information technology - Cloud computing - Reference architecture".

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] Italian Digital Agency: "Three-Year Plan for ICT in Public Administration (2017 - 2019)".

NOTE: Available at https://pianotriennale-ict.readthedocs.io/en/latest/.

[i.2] National Institute of Standards and Technology NIST SP 800-122: "Guide to Protecting the Confidentiality of Personally Identifiable Information (PII)".

[i.3] ETSI TS 133 401: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 3GPP System Architecture Evolution (SAE); Security architecture (3GPP TS 33.401)".

[i.4] 3GPP TR 22.864: "Feasibility study on new services and markets technology enablers for network operation; Stage 1".

[i.5] ISO/IEC 19944:2017: "Information technology - Cloud computing - Cloud services and devices: Data flow, data categories and data use".

[i.6] FP7-ICT 611659 AU2EU Deliverable D4.2.1: "Cryptographically enforced access control".

NOTE: Available at http://www.au2eu.eu/uploads/Publications/deliverables/AU2EU_D4.2.2_Final.pdf.

[i.7] 5G Ensure project: "Deliverable D2.1: Use Cases".

NOTE: Available at http://www.5gensure.eu/sites/default/files/Deliverables/5G-ENSURE_D2.1-UseCases.pdf.

[i.8] F. van den Broek, R. Verdult, J. de Ruiter: "Defeating IMSI Catchers".

NOTE: Available at http://www.cs.ru.nl/~rverdult/Defeating_IMSI_Catchers-CCS_2015.pdf.

[i.9] P. Paillier: "Public-Key Cryptosystems Based on Composite Degree Residuosity Classes", EUROCRYPT, pages 223-238. Springer, 1999.

NOTE: Available at https://link.springer.com/chapter/10.1007/3-540-48910-X_16.

[i.10] ETSI TR 101 567: "Lawful Interception (LI); Cloud/Virtual Services for Lawful Interception (LI) and Retained Data (RD)".

[i.11] L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, W. Jonker: "Information security applications", pages 309-323, Springer-Verlag, Berlin, Heidelberg, 2009.

[i.12] A. Sahai, B. Waters: "Fuzzy Identity Based Encryption", Advances in Cryptology - EUROCRYPT, Volume 3494 of LNCS, pages 457-473. Springer, 2005.

[i.13] V. Goyal, O. Pandey, A. Sahai, B. Waters: "Attribute-based encryption for fine-grained access control of encrypted data", Proceedings of the 13th ACM Conference on Computer and Communications Security, CCS '06, pages 8-98, New York, NY, USA, 2006. ACM.

[i.14] J. Bethencourt, A. Sahai, B. Waters: "Ciphertext-policy attribute-based encryption", Proceedings of the 2007 IEEE Symposium on Security and Privacy, SP'07, pages 32-334. Washington, DC, USA, IEEE Computer Society.

[i.15] A. Boldyreva, V. Goyal, V. Kumar: "Identity based encryption with efficient revocation", Conference on Computer and Communications Security, pages 417-416, 2008.

[i.16] ETSI TR 187 010: "Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Security; Report on issues related to security in identity imanagement and their resolution in the NGN".

[i.17] M. Piretti, P. Traynor, P. McDaniel, B. Waters: "Secure attribute-based systems", Journal of Computer Security, 18(5), pages 799-837, 2010.

[i.18] Z. Xu, K. Martin: "Dynamic User Revocation and Key Refreshing for Attribute-Based Encryption in Cloud Storage", Trust, Security and Privacy in Computing and Communications (TrustCom), IEEE 11th International Conference, 2012, pp. 844-849.

[i.19] Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[i.20] ISO/IEC 29100:2011: "Information technology - Security techniques - Privacy framework".

[i.21] ETSI TS 103 532: "CYBER; Attribute Based Encryption for Attribute Based Access Control".

[i.22] Directive 2012/18/EU of the European Parliament and of the Council of 4 July 2012 on the control of major-accident hazards involving dangerous substances, amending and subsequently repealing Council Directive 96/82/EC Text with EEA relevance.

[i.23] IEEE 802.11: "IEEE Standard for Information technology--Telecommunications and information exchange between systems Local and metropolitan area networks--Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," in IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012) , vol., no., pp.1-3534, Dec. 14 2016.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

cloud platform provider: cloud service provider providing identity management services and interfaces for third party applications using the platform services

cloud platform user: cloud service user consuming one or more platform services

cloud service customer: individual or organization consuming one or more cloud services provided by a Cloud Service Provider

cloud service partner: individual or organization providing support to the provisioning of cloud services by the Cloud Service Provider, or to the consumption of cloud service by the Cloud Service Customer

cloud service provider: individual or organization providing cloud services to one or more Cloud Service Customers

cloud service user: individual consuming one or more cloud services using a particular device

data subject: identified or identifiable natural person to which the data relates, or device that produces data that can be linked to a natural person

NOTE: In the sense of the GDPR [i.19], an identified or identifiable natural person to which the data relates. In the present document, this definition is extended to devices that produce data that can be linked to a natural person. See also PII principal.

direct access: access to data that is available in cleartext via a software-based access control system

generated data: data that is the result of an analytical process performed on behalf of, and which still relate to, the data subject

NOTE 1: Typically, generated data can be the result of a process applied to operational data.

NOTE 2: Depending on their characteristics, generated data can fall into the category of personal data as defined by the GDPR [i.19].

home network: central source for mobility services to the subscriber

NOTE: The subscriber has a direct subscription with the Home Network.

indirect access: access to data that is available in ciphertext form and requires the separate provisioning of a key followed by a decryption step, before the data can be accessed

key management: administration and use of the generation, registration, certification, deregistration, distribution, installation, storage, archiving, revocation, derivation and destruction of keying material in accordance with a security policy

location data: data composed of a position and a timestamp.

lone worker: employee who performs an activity that is carried out in isolation from other workers without close or direct supervision

operational data: data that is captured by a device on behalf of the data subject

NOTE 1: In an industrial context, this includes data captured by the worker's equipment as well as by facilities provided by the workplace (e.g. the physical access control system of a restricted area).

NOTE 2: Depending on their characteristics, operational data can fall into the category of personal data as defined by the GDPR [i.19].

personal data: any information relating to an identified or identifiable natural person ('Data Subject')

Personally Identifiable Information (PII): any information that (a) can be used to identify the PII principal to whom such information relates, or (b) is or might be directly or indirectly linked to a PII principal

NOTE 1: To determine whether a PII principal is identifiable, account can be taken of all the means which can reasonably be used by the privacy stakeholder holding the data, or by any other party, to identify that natural person [i.20].

NOTE 2: In the US, according to [i.2]: any information about an individual maintained by an agency, including any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information.

PII controller: privacy stakeholder that determines the purposes and means for processing personally identifiable information (PII) other than natural persons who use data for personal purposes [i.20]

PII principal: natural person to whom the personally identifiable information (PII) relates [i.20]

PII processor: privacy stakeholder that processes personally identifiable information (PII) on behalf of and in accordance with the instructions of a PII controller [i.20]

platform provider: service provider providing services necessary to support a platform

processing of PII: operation or set of operations performed upon personally identifiable information (PII) [i.20]

NOTE: Examples of processing operations of PII include, but are not limited to, the collection, storage, alteration, retrieval, consultation, disclosure, anonymization, pseudonymization, dissemination or otherwise making available, deletion or destruction of PII [i.20].

serving network: home network or visited network the user equipment is connected to

subscriber User Equipment (UE): any device allowing a user access to network services

static data: data that has been configured by the owner (e.g. an employee) on their devices, or by the device manager (e.g. the employer) on behalf of the owner

NOTE: Depending on their characteristics, Static Data can fall into the category of personal data as defined by the GDPR [i.19].

trust: level of confidence in the reliability and integrity of an entity to fulfil specific responsibilities

Trusted Authority (TA): ABE authority entitled to generate the master public key MPK and the corresponding secret keys according to a selected large universe ABE scheme

NOTE: A Trusted Authority can be implemented as a Key Management Server (KMS) and can be distributed across several servers residing on different domains.

untrusted storage: memory storage that is not trusted to provide adequate confidentiality and access control guarantees for the stored data

visited network: any network that interacts with the Home Network to provide mobility services to the subscriber terminal

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

2G	2nd Generation (mobile networks)
3G	3nd Generation (mobile networks)
4G	4nd Generation (mobile networks) also known as LTE

5G	5nd Generation (mobile networks)
ABAC	Attribute Based Access Control
ABE	Attribute Based Encryption
AKA	Authentication Key Agreement
ANPR	Anagrafe Nazionale della Popolazione Residente
AP	Access Point
API	Application Programming Interface
AV	Authentication Vector
CA	Certification Authority
CP-ABE	Ciphertext Policy Attribute Based Encryption
CPU	Central Processing Unit
CSPa	Cloud Service Partner
CT	CipherText
EU	European Union
GDPR	General Data Protection Regulation
GUTI	Globally Unique Temporary Identifier
IA	Issuing Authority
ICT	Information and Communication Technology
IMEI	International Mobile Station Equipment Identity
IMSI	International Mobile Subscriber Identity
IoT	Internet of Things
KMS	Key Management Server
KP-ABE	Key Policy Attribute Based Encryption
LEA	Law Enforcement Authority
LI	Lawful Interception
LTE	Long Term Evolution
MAC	Medium Access Control
MCC	Mobile Country Code
MME	Mobility Management Entity
MNC	Mobile Network Code
MNO	Mobile Network Operator
MPK	Master Public Key
MSIN	Mobile Subscription Identification Number
MSK	Master Secret Key
OTA	Over The Air
PII	Personally Identifiable Information
PK	Public Key
PKI	Public Key Infrastructure
PLMN	Public Land Mobile Network
PLMN-id	Public Land Mobile Network Identifier
PP	Platform Provider
PSK	Pre-Shared Secret
Pu	Platform user
RADIUS	Remote Authentication Dial-In User Service
SAML	Security Assertion Markup Language
SIM	Subscriber Identity Module
SK	Secret Key
SOA	Service Oriented Architecture
SOAP	Simple Object Access Protocol
SP	Service Provider
SR	Service Requestor
STA	Station
TA	Trusted Authority
TR	Technical Report
UE	User Equipment
WEP	Wired Equivalent Privacy
WLAN	Wireless Local Area Network
WM	WLAN Manager
WP	WLAN Provider
WPA	Wi-Fi Protected Access

4 Mobile use case

4.1 Introduction

4.1.1 Scenario

This solution presents an alternative mechanism to conceal permanent or long-term subscription identifier when a temporary subscription identifier is not available (c.f. initial Attach Request in LTE). Information provided over the air during initial attach request in LTE can allow a passive eavesdropper to obtain a variety of sensitive information including, but not limit to: user identity, user subscribed services such as voice, messaging, data et cetera, location information, traffic, network usage [i.4].

By design, current mobile networks need to occasionally expose long terms identities such as IMSI or IMEI, but in some circumstances (e.g. before network attach is complete), no protection can be offered in all current systems. This is because mobile telecommunication systems (e.g. 2G, 3G, 4G) mandate the use of symmetric-key cryptography schemes, with pre-shared secret key, for guaranteeing user/network authentication and confidentiality and integrity of data sent over the air. Such schemes are based on the assumption that the user has been previously identified by the network. In the AKA method (Authentication Key Agreement) [i.3] the user is required to first provide his/her subscriber identifier (i.e. the IMSI - International Mobile Subscriber Identity) in order for the mobile network to be able to subsequently retrieve the user's credentials (including the user-specific pre-shared secret key) and perform the user authentication procedure. The subscriber identifier is sent to the mobile network in clear text, since initially no cryptographic material is yet available/shared between the user and the network before the conclusion of the AKA procedure.

EXAMPLE: John is visiting a foreign country and switches on his mobile. The visited PLMN requests John's user identity (e.g. the IMSI) in order to authenticate him. John implicitly relies on the assumption that the visited PLMN is a trusted network. Unauthorized disclosure of sensitive PII can happen if interconnected networks are not "legitimate" as expected to be [i.7].

The lack of protection can lead to users' privacy breaches such as the disclosure of subscription identifier to an unauthorized party, the subscriber spoofing, and the detection of subscriber's presence in certain location. An attacker can gather, for example by means of an IMSI sniffer, all IMSIs that are active in a certain geographic area. An IMSI sniffer can achieve this in two different manners: passive and active. A passive sniffer will be simply observing unencrypted wireless traffic and storing all observed IMSIs. An active sniffer will be using a fake base station (fake BTS - Base Transceiver Station), to which mobile phones (UE) in the neighbourhood will attempt to connect due to the detection of a stronger radio signal, and the fake base station will request (e.g. with an Identity Request message) each UE to identify itself. The active IMSI sniffing is also known as IMSI catching in the mobile networks environment, and it is considered a feasible attack and, in recent years, even affordable by using out-of-the-shelves tools.

IMSI sniffing/catching attacks mostly relate to the issue of users' location privacy, as the transmission of IMSI reveals the user approximate location. Location privacy attacks attempt to link an identity to a location. User's location tracking is performed by sniffing the user identities sent in clear text over the air. An attacker can collect users' identities in an area or place (e.g. in an airport) and can further on track the users' presence and movements.

The information can also be used to impersonate the user or to cause denial of service.

4.1.2 Preliminary considerations

A major problem in current mobile networks is that identifiers are occasionally exposed in situations where no security context (i.e. shared cryptographic material) is available, neither to authenticate identity requests, nor to protect (encrypt) the IMSI in the identity response, as well as in attach request messages, thus enabling UEs tracking.

The security problem is briefly summarized below:

- The IMSI is transmitted in clear text in the first Attach Request.
- Identity Requests are not authenticated and are used to retrieve the long-term user identity, e.g. in the case when the temporary identity (e.g. GUTI in 4G) results as invalid. This implies that the user's Identity Response contains the IMSI in clear text.

- Encryption of signalling is required to transmit subscriber's identities in a protected way. However, availability of encryption depends on the network configuration.

Temporary identities reallocation depends entirely on operator configuration. This situation can be exploited by attackers as described in the literature [i.8].

In comparison with symmetric-key cryptography method like AKA [i.3], the use of asymmetric-key (or public-key) cryptography can provide an increased level of protection of sensitive data exchange during user identification. In principle, to initiate the network Attach procedure, the subscriber's identity could be encrypted using the public key of the network operator. This way, the UE would not need to send its identity in the clear.

In particular, public-key encryption can be needed in the following situations:

- Attach Request - contains IMSI in clear text: the IMSI (and maybe other sensitive information, e.g. network and security capability) should be encrypted with the public key of the serving/home network (the public key is stored on the SIM, while the private key is available on the 5G Mobile Management Entity (MME)). Randomized encryption schemes should be applied in order to prevent linkability. Therefore, the risk that IMSI catchers can read, guess or track the IMSI is reduced.

- Identity Response - contains IMSI in clear text: the IMSI should be encrypted with the public key of the network (the public key is stored on the SIM, while the private key is available on the 5G MME). Randomized encryption schemes can be applied, such as homomorphic cryptosystems (e.g. [i.9]) or ABE encryption schemes. Therefore the risk that IMSI catchers can read or guess the IMSI is reduced.

If traditional public-key encryption were used to avoid an unauthorized disclosure of its identity, the UE would have to use the public key of the visited serving network. This would be impractical, as the UE would have to be securely provisioned with all public keys (more precisely, with unexpired public-key certificates or certificate chains with trusted public keys in the corresponding digital signatures) of all possible networks it can roam into and select the one to use based on the PLMN (Public Land Mobile Network) identifier. A public-key certificate signed by an authority binds a public key to an identity (e.g. a network identity). Public Key Infrastructure (PKI) is needed to generate and manage the required public-key certificates and the public keys used for signing the certificates need to be all trusted.

The present clause is intended to tackle this problem using a technique based on attribute based encryption. To this end, the following high level requirements are defined.

4.2 High level requirements

The present document specifies the following high level requirements (table 4.1) for the Mobile use case.

Table 4.1: Requirements derived from the "Mobile" use case

ID	Requirement
MO#1	The solution shall protect the entire long-term identifier (MCC, MNC and MSIN in IMSI or IMEI) over the air and ensure full perfect anonymity, with respect to unauthorized entities (in complexity-theoretic setting). This implies that the encryption shall be randomized, i.e. each time the same long-term identifier is encrypted, the ciphertext is computed as purely random.
MO#2	The solution shall discourage the fake BTS attack scenarios.
MO#3	The solution shall not necessarily require a Public Key Infrastructure (PKI) to generate and manage public-key certificates.
MO#4	The solution shall be Lawful-Interception (LI) friendly and shall allow the satisfaction of LI requests, providing the ability of the serving network to identify all the mobile network services of the LI target without the home network's assistance or visibility, preferably without noticeable delay.
MO#5	The solution shall address backwards compatibility requirements, e.g. regarding the access of 5G terminals to previous 4G networks (this requirement may have a strong influence on the 5G security design and complexity, and so will requirements on mobility - seamless or not - between different generations of mobile systems).
MO#6	The solution shall be scalable: it shall use a single, global, master public key MPK for encryption and a multiplicity of secret keys for decryption in possession of authorized entities (e.g. mobile operators). The secret keys shall be generated from the MPK and a master secret key MSK.
MO#7	Both MPK and MSK shall be independent of the number of authorized entities.
MO#8	The UE shall be provisioned with only one public key MPK used to encrypt its subscription identifier, independently of the serving network to which it is connected at any given time. (This requirement provides confidentiality protection also in case of roaming scenarios. It also avoids the need for updating the UE in case where the subscription changes).
MO#9	The solution shall be flexible: secret keys for decryption may be generated initially or in real time, upon verification of the identity attributes of the authorized entities. Identity attributes shall be public. New attributes and secret keys for decryption may be generated and added to the system at any time.
MO#10	The solution shall support attribute revocation, in which case new attributes and secret keys shall be used. Revoked attributes shall not be used.

4.3 Use case

4.3.1 Stakeholders

A 3GPP mobile network architecture [i.3] is assumed. In particular, the following roles are defined:

- Home Network: is the central source for mobility services to the subscriber. The subscriber has a direct subscription with the Home Network: an identity for the subscriber has been allocated before any access to the network. The Home Network interacts with the visited networks.

- Visited Network: is any network that interacts with the Home Network to provide mobility services to the subscriber terminal. A subscriber has an indirect subscription with the Visited Network, i.e. a valid subscriber identity is only allocated as part of the first access to the Visited Network.

- Serving Network: is a Home Network or Visited Network the user equipment is connected to. Each Serving Network has assigned an identifier PLMN-id (Public Land Mobile Network Identifier) composed of a Mobile Network Code (MNC) and a Mobile Country Code (MCC).

- Subscriber User Equipment (UE): is any device allowing a user access to network services. The Subscribed User Equipment is provided with at least one subscriber Identity Module.

- Trusted Authority (TA): is a global entity performing key management functions, which can be implemented as a Key Management Server (KMS). In the setup, it generates the master secret key MSK and the global, master public key MPK. MPK is used for encryption. The secret keys, which are used for decryption, are generated from MSK and MPK according to a selected large universe ABE scheme.

NOTE 1: In a large universe ABE scheme (e.g. [i.13] and [i.17]) , the number of attributes is not predetermined and the Trusted Authority can issue secret keys for new attributes at any time, while maintaining the same global master public key MPK and master secret key MSK.

Each issued secret key is bound to the global public key MPK and to a public identity attribute of an authorized entity to whom the secret key is provisioned (e.g. PLMN-id of a mobile operator). It can be used for decryption if and only if the same identity attribute, together with the global MPK, is used in the encryption. Otherwise, the decryption would not work.

Distributed TA: by using secret sharing and threshold cryptography the TA can be distributed among a number of independent servers, belonging to different administrative domains (e.g. controlled by operators running Visited Networks) and placed at different geographic locations. (For cybersecurity, they should preferably be running on different operating systems.) Both security and reliability/robustness can be achieved by using a (k, n)-threshold secret sharing scheme. In such a scheme, a secret is distributed in terms of shares stored at individual servers, so that any k out of n servers can reconstruct the secret, whereas any k-1 or less servers do not get any information about the secret from their shares combined. This means that up to n-k servers are allowed to fail (robustness), while up to k-1 servers are allowed to be compromised (security). In the case of ABE, the secret is the secret master key MSK whose shares are independently generated and stored at individual servers.

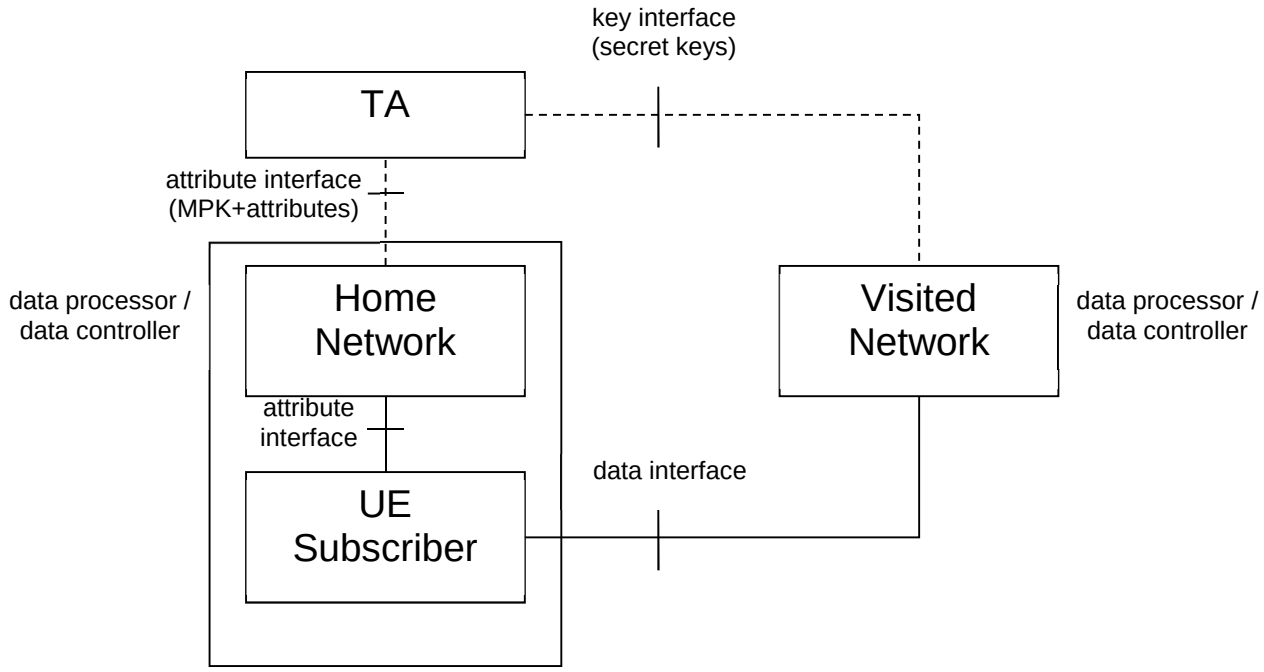

Figure 4.1: Architecture

NOTE 2: The distribution of the public key to the Subscriber UE and the distribution of the secret keys to the Visited Networks is out of scope for the present use case and given as a precondition (see clause 4.3.2).

4.3.2 Preconditions

Relying on a trusted hardware, the Trusted Authority (TA) securely generates the master secret key (MSK), the global master public key (MPK), and the corresponding secret keys according to a selected large universe ABE scheme. As the TA uses a large universe ABE scheme, secret keys can be generated initially or in real time, upon verification of the identity attributes of the authorized serving networks. The generated secret keys are bound to MPK and these identity attributes.

Identity attributes are public and specific to the Visited Networks that the UE is attaching to. The identity attribute of a Visited Network should have a standardized structure. It can include the network identifier (MNC, MCC) and a version or sequence number in order to support revocation. In the case of revocation, the new attribute and the corresponding new secret key start to be used, while the revoked attribute is included in the revocation list. Such a list can be occasionally updated and over-the-air provisioned to the UE by the home network to be stored there.

Alternatively, the identity attribute can include an expiry time (e.g. expiry date) of the respective secret key. The built-in expiry date enables occasional refreshing of the secret key and automatically renders an old secret key obsolete after the expiry date. Namely, the UE will then perform encryption only if the current date is consistent with the attribute expiry date received in broadcast.

Each Visited Network participating to the scheme is considered an authorized entity and is securely provisioned with a secret key corresponding to the attribute that the UE will use during the encryption.

NOTE: As in any public-key based encryption mechanism, the MPK (and consequently the MSK) can be occasionally (e.g. periodically) updated. In this case, the new MPK can be Over-The-Air (OTA) provisioned by each mobile operator to the respective subscriber's UE.

Each subscriber UE is securely provisioned (e.g. preloaded) with the global public key by its Mobile Network Operator (MNO). The UE uses this global public key independently from the network where the UE is attaching.

In a distributed TA, independently generated shares of the MSK are stored in individual servers and MSK itself is neither stored nor reconstructed at any time. A homomorphic property of the secret key generation function, with respect to MSK and a randomized parameter used in the secret key generation (e.g. as in [i.13] and [i.17]), enables threshold cryptography. It means that the secret key can be computed securely in a distributed way by each server computing a partial function on its share of MSK and its share of the randomizing parameter, and by combining the partial computation results of the servers by the network operator.

4.3.3 Trigger

The UE tries an Initial Attach request to the Visited Network.

4.3.4 Flow of events

The Visited Network broadcasts (continuously) its assigned attributes together with its network identifiers (MCC and MNC).

The UE encrypts the subscription identifier (IMSI) according to an ABE scheme by using the global public key MPK and the attribute received in broadcast over the air (the attribute identifying the Visited Network to which the UE is currently connected to) and transmits the encrypted value to the (Mobile Management Entity of the) Visited Network by an Attach Request. Randomized encryption in ABE enables the bit size of encrypted IMSI to be the same as the bit size of IMSI.

The Visited Network decrypts the received ciphertext by using its own secret key and obtains the IMSI.

NOTE: Upon decryption, the subscription identifier is known only to the authorized entity, i.e. to the current Visited Network. The Visited Network can then immediately start preparing lawful interception without the Home Network's assistance or visibility. In parallel, it forwards the subscription identifier to the Home Network, so that the lawful interception can start immediately after receiving back the AV from the Home Network.

The Visited Network forwards the IMSI to the UE's Home Network and requests the authentication data (AV).

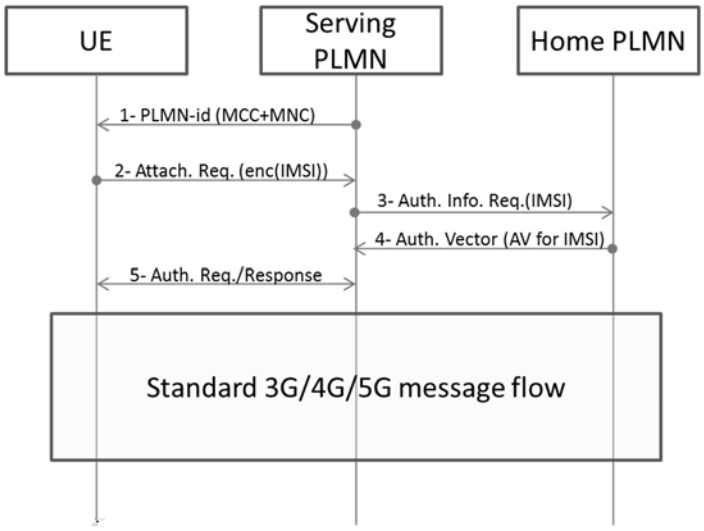

Figure 4.2: Message flow

4.3.5 Exit Condition

The Authentication Vector is returned to the Visited Network.

The standard 3G/4G/5G authentication procedure between the UE and the Visited Network is finally executed.

4.3.6 Security Aspects

In the 3G/4G/5G system, the network identifiers (MCC, MNC) of the participating network operators are assumed to be publicly known. The trust framework is regulated by inter-operator roaming agreements, and the involved operators are trusted to function accordingly, because of their business interests. In the case of ABE, the used network attributes include additional version or sequence numbers or expiry times of the secret keys corresponding to the network attributes. The network attributes are publicly known and managed by the TA, which can be made trustworthy by using standardized secure hardware components and by being implemented in a distributed way explained above. More generally, the needed trust can be further reduced by using multiple TAs, whose public index then needs to be included in the corresponding network attribute, in order for the UE to recognize which MPK to use for the encryption.

The TA securely generates and provisions the secret keys with embedded network attributes to the registered network operators, upon verifying the attributes. A secret key can perform decryption if and only if the subscription identifier is encrypted by the global public key MPK and the corresponding network attribute, which is received in broadcast by the UE. Before encryption, the UE can also check the locally stored list of operators having a roaming agreement with the home network. Accordingly, fake entities which are not in possession of a valid network attribute and a valid secret key corresponding to this attribute cannot perform decryption.

Consequently, a passive IMSI catcher can be established only by an entity in possession of a compromised secret key of the currently Visited Network, whereas an active IMSI catcher can possibly be established by an entity in possession of a compromised secret key of any registered network operator (which may depend on the roaming agreements relevant for a given area). Of course, any encryption method fails if the decryption key is compromised. However, in the case of IMSI protection, IMSI is made available to any Visited Network for lawful interception anyway. In this sense, the network operators need to be trusted unavoidably. So, an IMSI itself cannot be considered as confidential, but only that the UE with this IMSI is present at a given location at a given time. Moreover, in the 3G/4G system there is no IMSI protection at all.

4.3.7 Recommended ABE scheme

The CCA-secure KP-ABE with KP-FAME-KEM [i.21] ABE scheme should be used.

5 Privacy-Preserving federated WLANs use case

5.1 Introduction

5.1.1 Scenario

Wireless Local Area Networks (WLANs) are increasingly ubiquitous: with the growing number of smart mobile devices, users are willing to have access to the Internet everywhere. A solution for crowded cellular networks is to offload traffic, typically to WLANs. In this scenario, federated WLANs allow mobile users to access different WLANs in different places using the same credentials. However, traditional federated WLANs do not preserve the privacy of the users and require the maintenance of an online infrastructure to verify the credentials of the users.

This scenario addresses privacy-preserving federated WLANs. In this scenario, users can access different WLAN networks using their credentials (issued by different authorities/domains) while preserving their privacy. Federated WLAN providers receive guarantees that users are authorized to access the WLANs without the need for an online verification infrastructure.

The WLAN federation is composed by WLAN providers and authorities. WLAN providers manage a set of Access Points (APs) to which Stations (STA, i.e. user mobile devices) connect. APs announce their presence by broadcasting management packets called beacons. Authorities can issue credentials to users based on their attributes. WLAN providers can define WLAN access policies based on these attributes.

5.1.2 Preliminary considerations

Traditional WLAN network access relies on the following methods:

1) open (i.e. unencrypted) WLAN network in which typically the user is redirected to a splash-page for authentication;

2) WPA2-PSK secured WLAN network, in which all the legitimate users know the pre-shared secret (PSK). Security hence relies on the fact that this PSK is not leaked by end users, and is not disclosed via dictionary or brute-force attacks;

3) 802.1X-based secure WLANs, which require an ad-hoc setup on the user device and employ online interactive authorization using an authentication service leveraging a backend online infrastructure (e.g. RADIUS servers/proxies) which identifies the user.

The present document does not consider WLANs relying on other insecure protocols such as WEP and WPA.

Beacons generated by Access Points, containing the name of the network and other information, can be extended.

WPA2-PSK secrets are used to obtain a session key. This implies that if the WPA2-PSK of an AP changes, the users which are already connected to the AP do not need to be disconnected.

5.2 High level requirements

The present document specifies the following high level requirements (table 5.1) for federated WLANs use case.

Table 5.1: Requirements derived from the "Federated WLANs" use case

ID	Requirement
FW#1	The solution shall support an authentication mechanism preserving the privacy of the users. In particular, the federated WLAN infrastructure managers shall not be able to trace the user across different WLANs or on subsequent connections to the same WLAN.
FW#2	The solution shall be compatible with the legacy WLAN infrastructure.
FW#3	The solution shall allow the definition of access policies encompassing attributes belonging to different domains and shall guarantee that only authorized users can access the WLANs.
FW#4	The solution shall allow the issuance of credentials to users in form of secret keys.
FW#5	Credentials may be issued from different domains (e.g. companies, municipalities, universities, etc.).
FW#6	The solution shall be flexible: new attributes and secret keys may be generated and added to the system at any time.
FW#7	The solution shall support attribute revocation, in which case new attributes and secret keys shall be used. Revoked attributes shall not be used.

5.3 Use case

5.3.1 Stakeholders

An IEEE 802.11 [i.23] architecture is assumed (see figure 5.1), with additional actors and roles. In particular, the following roles are defined:

- Station (STA): is the mobile user device which connects to an Access Point.
- Access Point (AP): is managed by a WLAN provider. One or more STA can connect to the same AP.
- Issuing Authority (IA): is a CP-ABE authority entitled to generate the public key MPK and the corresponding secret keys according to a selected large universe ABE scheme. Multiple IA entities can coexist in the same system.
- WLAN Manager (WM): is used by the WLAN Provider (WP) to define policies on a set of APs.

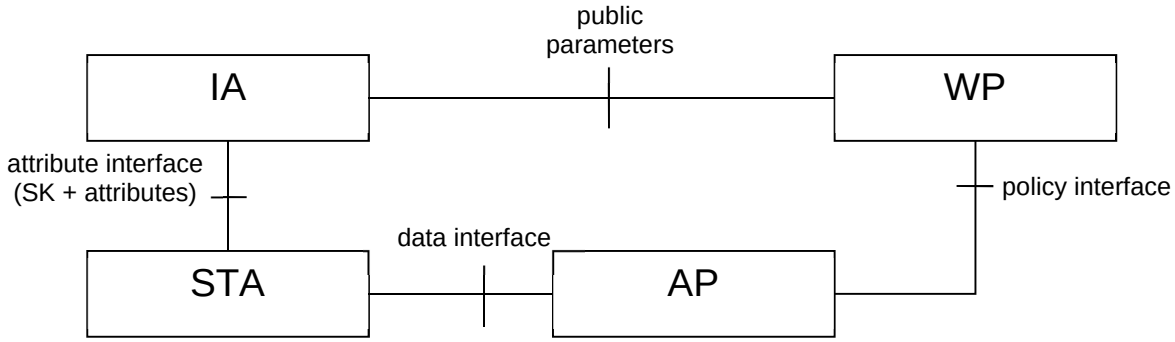

Figure 5.1: Architecture

5.3.2 Preconditions

The WP has received the public ABE parameters from the IAs.

The WP has defined an access policy.

The access policy alongside with the public ABE parameters have been transferred to the AP.

The AP has generated a random WPA2-PSK and has encrypted it using the provided policy.

The STA has received ABE secret keys from the IAs based on the user attributes.

5.3.3 Trigger

A user instructs a STA to connect to an AP.

5.3.4 Flow of events

The AP broadcasts chunks of the encrypted WPA2-PSK secret in the generated IEEE 802.11 beacons.

The STA collects the chunks in the beacons and reconstructs the encrypted WPA2-PSK secret.

If the attributes in the ABE secret keys satisfy the access policy of the AP, the STA can decrypt the WPA2-PSK secret.

The STA randomizes its MAC address and uses the WPA2-PSK secret to connect to the AP.

5.3.5 Exit condition

The STA connects successfully to the AP.

After this event, the AP generates a new random WPA2-PSK secret and encrypts it according to the access policy.

5.3.6 Recommended ABE scheme

The CCA-secure CP-ABE with CP-FAME-KEM [i.21] ABE scheme should be used.

6 Internet of Things use cases

6.1 Overview

The present clause describes several use cases related to the Internet of Things and progressively illustrates the capabilities to be supported by a cryptosystem providing access control.

In a first use case, the use of asymmetric encryption in order to export data to untrusted storage is introduced. Then, scenarios where access control should apply to data in ciphertext form are presented. In a third use case, the point is made that the definitions of new access control policies should be possible for existing ciphertext. Similarly, existing policies should be applicable to data that is generated from existing ciphertext. In a fifth use case, constraints related to offline access control are presented. In a sixth use case, it is reminded that not all data may exist in ciphertext form, and thus the concepts of direct and indirect access are introduced. In a seventh use case, concrete examples of access control in the industrial domain are provided.

High-level requirements for each use case are summarized in clause 6.2.

6.2 High level requirements

The present document specifies the following high level requirements (table 6.1) for the "Securing and exporting data to untrusted storage" use case (clause 6.3.1):

Table 6.1: Requirements derived from "Securing and exporting data to untrusted storage"

ID	Requirement	Comment
IoT#1	The solution shall be designed in such a way that a device can generate ciphertext but may not necessarily be able to recover the plaintext.	This is possible with an asymmetric cryptosystem where the keys to encrypt and decrypt are not the same.

The present document specifies the following high level requirements (table 6.2) for the "Bundling encrypted data with access control capabilities for use in an industrial context" use case (clause 6.3.2):

Table 6.2: Requirements derived from "Bundling encrypted data with access control capabilities for use in an industrial context"

ID	Requirement	Comment
IoT#2	The solution shall allow that access control policies defined for data in plaintext are applicable to the same data in ciphertext form.	
IoT#3	The solution shall allow for a variety of subjects (in the access control sense).	Scenario #1
IoT#4	The solution shall allow that only one ciphertext corresponding to a protected datum is necessary in order to support all access control policies that may apply to the datum.	Scenario #2
IoT#5	The solution should allow for devices and functions to be subjects (in the access control sense).	Scenario #2
IoT#6	The solution should allow for leveraging an access control policy in order to perform identification or authentication of users and devices.	Scenarios #2 & #3, as well as offline case in clause 6.3.6
IoT#7	The solution shall support addition and revocation of data subjects (in the access control sense).	
IoT#8	The solution shall support changes to access rights granted to subjects, as well as automatic expiration of those.	

The present document specifies the following high level requirements (table 6.3) for the "Assigning new access control policies to already encrypted data" use case (clause 6.3.3).

Table 6.3: Requirements derived from "Assigning new access control policies to already encrypted data"

ID	Requirement	Comment
IoT#9	The solution shall support update and expiration of access control policies, including for data that is protected in ciphertext form.	Also expressed in clause 6.3.5.
IoT#10	The solution shall support data minimization, including for ciphertext that is provided as a response to a request.	This can be easily achieved in KP-ABE, when the ciphertext structure matches existing attributes
IoT#11	The solution shall support addition of new access control policy for existing ciphertext.	
IoT#12	The solution shall support that an access control policy is not based on the identity of the subject.	This stems from the last sentence in the use case. The subject might be identified when the access control policy is issued, but no identity can be embedded in the ciphertext.

The present document specifies the following high level requirements (table 6.4) for the "Applicability of access policies to processed data" use case (clause 6.3.4):

Table 6.4: Requirements derived from "Applicability of access policies to processed data"

ID	Requirement	Comment
IoT#13	The solution shall allow for a third-party to return generated data in a ciphertext form for which access control policies defined by the data subject apply.	Typically, in ABE, the third-party is given the master public key MPK.

The present document specifies the following high level requirements (table 6.5) for the "Offline access control in constrained operational environments" use case (clause 6.3.5):

Table 6.5: Requirements derived from "Offline access control in constrained operational environments"

ID	Requirement	Comment
IoT#14	The solution shall support out-of-band provisioning of access control policies, even when the data subject, or their device, is not accessible online.	
IoT#15	When existing access policy material is reused among several access control subjects, the solution should provide means to ensure traceability of access on a per-subject basis.	
IoT#16	The solution shall support that, within an existing system, devices are provisioned with new access control policies only when they need to access data from new subjects.	The reverse is supported by the first requirement in this table.
IoT#17	The solution shall support that a data subject issues an access control policy, even in the absence of a connection to the network infrastructure.	This can be viewed as a variant of the platform provider model.
IoT#18	In addition to supporting offline operations, the solution shall support situations where latency is a key performance indicator.	This has consequences on ABE schemes that are suitable for the Industrial IoT.

The present document specifies the following high level requirements (table 6.6) for the "Direct and indirect data access" use case (clause 6.3.6).

Table 6.6: Requirements derived from "Direct and indirect data access"

ID	Requirement	Comment
IoT#19	The solution shall support the definition of access control policies applicable both to the direct access and to the indirect access cases.	The Intermediate Access Control Language [i.21] can fulfil this requirement in addition to providing an adaptation layer between ABE and existing software-based access control systems.

The present document specifies the following high level requirements (table 6.7) for the "Access control examples in the Industrial Internet of Things" use case (clause 6.3.7).

Table 6.7: Requirements derived from "Access control examples in the Industrial Internet of Things"

ID	Requirement	Comment
IoT#20	The solution shall support time-based access control, whereby operational data is captured, or generated data is released, over a specific period of time.	
IoT#21	The solution shall support that an access control policy is disabled in exceptional circumstances.	See time-based access control for the industrial wearable.
IoT#22	The solution shall support position-based access control, whereby the position can be expressed in terms of coordinates or labels, and can be relative to another object or absolute in a cartographic system.	
IoT#23	The solution shall support access control based on the origin of the data.	These two requirements can greatly benefit from existing ontologies.
IoT#24	The solution shall support access control based on the data type. The data type hierarchy shall allow for supporting general cases and more specialized cases.	
IoT#25	The solution shall support mandatory access control in the form of an expected clearance level.	This is straightforward with CP-ABE. In KP-ABE, attributes can be reserved e.g. as clearance or role identifiers, in which case it is not an attribute that describes the data.
IoT#26	The solution shall support role-based access control.	
IoT#27	The solution shall support access control based on expected subject characteristics.	
IoT#28	The solution shall support access control based on service tiers.	
IoT#29	The solution shall support access control based on emergency levels.	

The present document specifies the following additional high level requirements (table 6.8), derived from previous requirements:

Table 6.8: Further requirements

ID	Requirement	Comment
IoT#30	The solution shall support attribute management.	This is a logical consequence of requirement #13, since the data processor may need to define new attributes to better characterize the generated data. In addition, clause 6.3.7 illustrates the variety of access control scenarios and hints that new scenarios may appear.
IoT#31	The solution shall allow for the identification and possibly authentication of third-parties contributing data for which access control policies defined by the data subject apply.	Natural consequences of requirement #13. This may not be supported by the ABE scheme itself but rather by having locations in the ciphertext structure, for example dedicated to digital signatures.
IoT#32	The solution shall allow for proving the integrity of the data protected in ciphertext form.	
IoT#33	The solution shall allow access control subjects to identify the data subject for which the policy is valid.	In a multi-party environment, the subject needs to identify the cryptosystem that was used to produce the ciphertext as well as the data subject, so that they can select the appropriate secret key.
IoT#34	The ciphertext structure shall allow for data granularity.	Continuation of requirement #10. Once the data is encrypted, the device has no means to select a subset of the data from the ciphertext (before handing it to a requesting third-party), other than the structure of the ciphertext itself and embedded metadata.
IoT#35	Methods should be available to ensure that ciphertext metadata does not constitute personal data in the sense of the GDPR [i.19], or does not allow to infer personal data about the data subject.	Non-ABE techniques (such as more traditional confidentiality protection mechanisms and anonymization) may be used. ABE attributes in a large universe can be pseudonymised.

6.3 Use cases

6.3.1 Securing and exporting data to untrusted storage

6.3.1.1 General use case description

The present use case illustrates a device that uses irreversible (from the point of view of the device) encryption in order to ensure a high level of confidentiality assurance of Operational Data and Static Data of a Data Subject. The Untrusted Storage can be local or remote. By doing so, the device can achieve higher confidentiality guarantees compared to runtime countermeasures.

6.3.1.2 Stakeholders

User: the user can be a data subject (if the data in question refers to the user) or not. In this scenario the user is a passive stakeholder which benefits from the functionality.

Device manager: the device manager can coincide with the user or be a different stakeholder responsible for the cryptosystem on the device.

6.3.1.3 Scenario(s)

When data is located on a user device, technical countermeasures can be used to address threats against the confidentiality of the data. This includes the use of a trusted execution environment (within a secure element or a hardware enclave) providing memory isolation, full-disk encryption providing data-at-rest protection, class-based file encryption, software-based access control, and so forth.

The choice of countermeasures depends on the risk analysis. Two contributing factors in the Industrial IoT case are operations in a multi-tenant environment (e.g. the device hosts third-party applications), and the physical security of devices in the field.

Most security countermeasures rely on the existence of secure memory which is expensive and available in limited quantity. Thus, the amount of data that a device can store with appropriate security guarantees can be very limited. Furthermore, there is always a risk that runtime countermeasures fail to prevent malicious extraction of the data before it is transferred away from the device.

In order to address this, the device encrypts and then transfers data to a memory location that is outside of the device secure environment. Such location can be a regular memory location on the device (for instance an eMMC memory chip or a microSD card), a physical medium, or a data storage service in the infrastructure (for instance a cloud area). In some cases, the device can relinquish access to the data upon encryption. When the encryption scheme is constructed such that the device never possesses the key material necessary to perform decryption, the level of confidentiality assurance of the protected data on the device is that of the encryption scheme and can be very high.

6.3.1.4 Information Flows

Pre-conditions: operational data and/or static data (hereunder referred to as data) reside in secure memory.

Post-conditions:

- Success: the data is stored in encrypted form in untrusted storage, and the secure memory is freed.
- Failure: the data remains in secure memory. Further data capture or generation can be impossible, resulting in degraded mode or the inability of the device to continue operations.

Begin triggers: the device detects that data is filling secure memory (e.g. by means of a quota), or a periodic task is scheduled.

Termination triggers: completion of the encryption and transfer to untrusted storage.

Working assumptions: the device is provisioned, by the device manager, with a public key suitable for the encryption step.

Normal flow:

- The device encrypts the data in the device secure environment.
- The encrypted data is transferred to untrusted storage.
- The corresponding plaintext data that resides in secure memory is erased.

Alternative flow(s): the data is securely transferred to a trusted third-party that will perform the data encryption.

6.3.1.5 Operational constraints

For the normal flow, efficiency constraints exist on the encryption scheme, in particular processing effort (CPU cycles and power consumption), processing memory footprint, ciphertext size and public key size, so that the encryption algorithm can be executed by an embedded device with limited capabilities.

6.3.2 Bundling encrypted data with access control capabilities for use in an industrial context

6.3.2.1 General use case description

This use case illustrates that once data has been encrypted, the enforcement of access control policies remains so that legitimate subjects can access the data. In a variant, access control is leveraged in order to perform device and user authentication.

6.3.2.2 Stakeholders

User: the user can be a data subject (if the data in question refers to the user) or not. The user can configure access control policies for the data.

Device manager: the device manager can coincide with the user or be a different stakeholder.

6.3.2.3 Scenario(s)

Once data has been encrypted as described in clause 6.3.1, the access control policies that have been defined for the data (static data, operational data, and generated data) in plaintext also apply to data in ciphertext form and are handled by the cryptosystem itself when the ciphertext is requested by a third-party. This approach can also be used for static data that the device does not need to access, but has been configured by the data subject. To illustrate this, some scenarios of the industrial domain are presented below.

In a first scenario, a worker is conducting monitoring of an equipment using a handheld device (equipped e.g. with near-field communications capabilities) or a remote console. The monitoring can be routine or done in response to an alarm. The worker is an employee of the factory owner, or a contractor from a service provider. They only have access to monitored data on a need-to-know basis, because the data can constitute personal data of other workers, or more generally confidential data of various stakeholders involved in the industrial plant.

In a second scenario, the capabilities of the access control system are leveraged to support two or more service functions that need to exchange data as part of an industrial process. These functions reside on different devices and there is a multiplicity of devices involved in the industrial process. However, device-to-device and function-to-function data exchange is restricted on a need-to-know basis to fulfil safety and security requirements. In order to limit bandwidth and storage consumption, the cryptosystem is designed so that a single ciphertext exists per protected datum, that supports access control policies. As the ciphertext is not necessarily located on the originating device any longer, not having to contact said device to decrypt the ciphertext further reduces bandwidth consumption.

In a third scenario, devices in an industrial plant are configured to monitor the worker in order to comply with safety best practices and regulations. This can take the form of a query over a wireless bearer, for example against an identifier stored in a wearable of the worker. However, not all devices in the industrial plant can query the identity wearable and obtain the worker identity (or other personal data such as health conditions, education background, trainings, certifications…), but only those that are trusted for performing the monitoring task and processing related data (e.g. storage of the identifier along with location) in compliance with data privacy regulations.

In a fourth scenario, the capabilities of the access control system are leveraged to provide simplified access control, such as access to a service function or a configuration panel. In this approach, the cryptosystem supports that an access control policy is enough to run an authentication protocol. This capability can be used for device as well as user authentication.

Considerations related to traditional access control apply to the cases above, in particular:

- At some given point, a new subject (user, device or function) is integrated in the system, or revoked from it.
- Access rights granted to subject can change over time (e.g. due to changes in responsibilities) and can be set to expire automatically after some time (e.g. for contractors).

6.3.2.4 Information Flows

Pre-conditions: there is data residing in secure memory. The system is configured with access control policies for the data.

Post-conditions:

- Success: access control policies are applicable to ciphertext, which can only be decrypted with a key that matches an access control policy applicable to the plaintext.
- Failure: the device is unable to encrypt the data according to the configuration of the access control system.

Begin triggers: data in secure memory is about to be encrypted and exported to untrusted storage as per the use case in clause 6.3.1.

Termination triggers: completion of the encryption and transfer to untrusted storage.

Working assumptions: the device is provisioned, by the device manager, with a cryptosystem supporting access control and a public key suitable for the encryption step.

Normal flow:

- The device retrieves access control information pertaining to the data that is about to be encrypted (e.g. the data attributes or the access control policies, depending on the cryptosystem variant).
- From the information gathered in the previous step, the device derives the appropriate cryptosystem parameters.
- The device encrypts the data according to the parameters in the previous step.

6.3.3 Assigning new access control policies to already encrypted data

6.3.3.1 General use case description

The present use case illustrates the allocation of new access control policies in order to allow a data consumer access to operational data that is already in ciphertext form. Examples are given with the management of location data.

6.3.3.2 Stakeholders

User: the user is a data subject.

Device manager: the device manager can coincide with the user or be a different stakeholder.

Data consumer: the consumer can be a data processor or not.

6.3.3.3 Scenario(s)

Due to their sensitive nature, personal data in the sense of the GDPR [i.19] benefit from privacy-by-design/privacy-by-default measures. Retention and disclosure of personal data to a third-party is governed by strict legal and contractual rules that protect the privacy of the data subject. This includes in particular the application of the "need to know" and "data minimization" principles. In terms of access control, the consequence is that the initial set of access policies is limited and a new access policy on an object (personal datum) is granted to a subject only when there is a justification for doing so. Access control policies are bound in scope, regularly audited to determine their relevance, and eventually updated or expired. When the operational data is stored in ciphertext form, the cryptosystem supports the management of access control policies, in particular addition and revocation.

The location of the data subject is one example of operational data that is also personal data. Minimization of location data can be achieved by restricting the time and position of the location data to periods and zones that are only relevant to the data request. Examples of cases when access to location data is granted after it has been stored in ciphertext form are given below:

- **Safety incident audit:** the worker is involved (as victim, witness or responsible party) in an industrial incident (e.g. an accident leading to an injury or loss of equipment). Location data relevant to the incident is shared with the auditor to identify the causes of the incident. In another variant of this case, the worker has developed health problems and the location data is used in order to determine whether there is a correlation between work zones and health issues.

- **Law enforcement:** the worker has been victim of a physical aggression. The location data relevant to the aggression is shared with law enforcement in support of an investigation. In other variants of this case, the worker can be witness, or suspected, of unlawful activities.
- **Insurance:** the worker's employer has contracted a new insurance policy covering industrial risks at the workplace. The insurance company requests location data relevant to working hours in the previous semesters, in order to evaluate the level of risk to which the worker is exposed. Such sharing scheme can also support the emerging on-demand insurance business model, where insurers ensure that a certain number of conditions have been respected and/or will be respected for an employer to benefit from a discounted premium insurance.
- **Social benefits:** in some countries, specific working situations are recognized as special for the calculation of social benefits (e.g. retirement and compensation of overtime). The time spent on duty, as well as the worker's location while on duty, can be used to identify these special working situations.
- **Business process improvement:** the location data is shared with an internal commission of the worker's company for the improvement of safety and performance. For example, to limit the worker's presence in dangerous areas, or to limit the distance that the worker covers daily.

In all the cases above it is possible that the access control subject is identified only at the time the data is to be shared with them.

6.3.3.4 Information Flows

Pre-conditions: there exists Operational Data in ciphertext form, pertaining to the data subject.

Post-conditions:

- Success: the request for accessing the Operational Data is evaluated. When it is rejected, no access control policy is provided and the processor cannot access the data. When it is accepted, an access control policy is generated and the processor can access the data.
- Failure: N/A.

Begin triggers: an existing or new actor (access control subject) requires access to Operational Data.

Termination triggers: N/A.

Working assumptions: the cryptosystem supports the enrolment of new access control subjects (e.g. processors) without modifying the ciphertext generated on behalf of the data subject. The system allows the data subject to manage access control policies (creation, deletion, expiration), and/or allows a third party to manage access control policies on behalf of the subject.

Normal flow:

- Operational Data of a data subject is stored in ciphertext form within the system.
- A processor requests access to a dataset.
- A procedure between the data subject (or a proxy thereof) and the processor takes place in order to evaluate the request (this can involve an authentication step as well as a non-automated step such as an evaluation by a commission).
- When the request is rejected, the requestor is informed that no access control policy is issued.
- When the request is accepted, a new access control policy is generated and the processor can access the dataset.

6.3.4 Applicability of access policies to processed data

6.3.4.1 General use case description

The present use case illustrates the situation where data is processed by a third-party (the processor) into a new dataset on behalf of a data subject. The data subject can issue access control policies on the generated data.

6.3.4.2 Stakeholders

User: the user is a data subject.

Device manager: the device manager can coincide with the user or be a different stakeholder.

Data consumer: the consumer is a data processor.

6.3.4.3 Scenarios

One purpose of data collection is to process the data into information (such as statistics) and new datasets. The uses of location information described in clause 6.3.3 are examples of such data processing. In some cases, the data subject retains control over the result of the processed data.

For example, by using location information, a PII processor is generating sport statistics on behalf of the data subject. By virtue of the contractual agreement between the data subject and the PII controller (or because of a legal obligation), the PII processor is not allowed to share the generated data with any other party. The PII controller and PII processor are also responsible for preventing data breaches when processing the data handled by the data subject.

Situations similar to the one described above exist in the industrial domain, where a variety of third-parties can provide data processing services with confidentiality constraints, in order to support an industrial process.

The data processor retrieves the data to process in ciphertext form, and the data subject or the device manager issues an adequate access policy for the processor. The cryptosystem supports that the processor returns the generated data in a ciphertext form allowing the data subject to issue access control policies on said generated data. The data subject thus retains control over the generated data. Another advantage of the procedure is that confidentiality of the data is guaranteed regardless of the security mechanisms available at the transport layers. Data is processed and generated in plaintext form only in the control domain of the data processor.

6.3.4.4 Information Flows

Pre-conditions: a cryptosystem for access control is configured for a set of devices, users and policies.

Post-conditions:

- Minimal success: the data subject retains control over the collected data.
- Success: the data subject retains control over the collected and generated data.
- Failure: N/A.

Begin triggers: a data processor collects data from the device.

Termination triggers: N/A.

Working assumptions: cryptosystem parameters, (access control policies, valid attributes) are pre-provisioned on devices.

Normal flow: Data subject retaining control over the collected data:

- The device encrypts the requested data according to defined policies and/or attributes.
- The ciphertext is returned to the processor.
- By matching the defined policy, the processor is able to decrypt the ciphertext and access the plaintext.
- The processor successfully uses the plaintext to generate data.

Alternative flow 1: Data subject retaining control over the generated data:

- The device encrypts the requested data according to defined policies; and/or
- The ciphertext is returned to the processor.
- By matching the defined policy, the processor is able to decrypt the ciphertext and access the plaintext.

- The processor successfully uses the plaintext to generate data.
- The processor allows the data subject to issue access control policies on said generated data, or identifies which attributes to use in order to annotate the generated data.
- The processor encrypts the generated data according to defined policies and/or attributes and returns it in a ciphertext form.

6.3.5 Offline access control in constrained operational environments

6.3.5.1 General use case description

The present use case illustrates the application of access control between two devices in an industrial context, without support of a network infrastructure and with minimal performance and provisioning overhead.

6.3.5.2 Stakeholders

Users: the users can be data subject or not.

Device manager: the device manager can coincide with the user or be a different stakeholder.

6.3.5.3 Scenario(s)

In the Industrial IoT domain, data caps and costs on the communication links between the device and the infrastructure strongly limits the use of the communication link for resolution of requests that involve a centralized infrastructure. This includes access control requests as well as management of access control policies. Furthermore, a communication link to the network infrastructure may not be available due to security restrictions or specific situations in the field (unavailability of radio coverage), or may not provide low enough latency to support fast decision-making process and/or data access requirement (for instance in case of critical issues for worker or machine requiring immediate actions).

The unavailability of a central authority that supports access control decisions and policy management would normally make the maintenance of the access control system difficult, especially in multi-stakeholder scenarios.

However, the cryptosystem is designed in such a way that it integrates out-of-band provisioning in order to support the functionalities described in clause 6.3.2.

In a first variant, existing access policy material is handed over to access control subjects and/or their devices and operates in a way similar to group keys. In this approach, additional means are necessary to allow for traceability of access on a per-subject basis.

In a second variant, the provisioning of the access control system is decentralized with each stakeholder in the industrial process providing on-demand, off-site pre-provisioning service related to the devices and equipment they are responsible for. Prior to operating on-site, an access control subject (e.g. a contractor) contacts each device manager in order to obtain access policies or attributes for relevant devices. This procedure between the access control subject and device managers can happen online when said subject is preparing the on-site visit, and supports that there is no communication channel available between the device managers and devices holding data. Expiration methods are in place to ensure that a subject cannot abuse the granted policies beyond their intended purpose.

The cryptosystem supporting authentication and access control is further optimized in such a way that the provisioning overhead is kept to a minimum. In particular, existing devices are only provisioned if they need access to data that is under the control of the newcomer device or user - so that there is no need to alter the configuration of existing devices in order to support access from the newcomer.

In this later situation, a newcomer device or user can issue an access control policy for existing devices even though there is no connection to the network infrastructure.

6.3.5.4 Information Flows

Pre-conditions: a cryptosystem for access control is configured for a set of devices, users and policies.

Post-conditions:

- Success: a new subject (device or user) is successfully integrated in the access control system without interaction of the network. Access control operates as expected.
- Failure: the new subject is not integrated in the access control system, or a misconfiguration prevents access control from operating properly.

Begin triggers: a new subject is set to be registered within the access control system.

Termination triggers: N/A.

Working assumptions: access control policies are pre-provisioned on devices.

Normal flow: The normal flow is described as follows:

- The new access control subject contacts the device manager to obtain access rights to data on relevant devices.
- The device manager provides the access control subject with appropriate cryptographic keys such that the access control subject can access the data.
- If necessary, the device manager updates cryptographic keys in existing devices that need access to data that is under the control of the newcomer access control subject.
- The device manager associates cryptographic keys with expiration methods to ensure that no subject can abuse the granted policies beyond their intended purpose.
- The new subject is successfully integrated in the access control system without interaction of the network.

6.3.5.5 Operational constraints

Use of a data link to the infrastructure is heavily restricted, if existent at all: an offline mode of operation is expected.

Communication latencies between devices in an industrial context are kept to a minimum in order to match operational constraints of the industrial process (in particular to avoid incidents).

6.3.6 Direct and indirect data access

6.3.6.1 General use case description

The present use case illustrates a situation where data to be accessed exists partly in plaintext and partly in ciphertext form. However, the access control policies that have been defined by the data subject are applicable to both the plaintext and ciphertext. Thus, the system can properly enforce access control in both cases. Direct and indirect data access are explained below.

6.3.6.2 Stakeholders

Data consumer: the consumer can be a data processor or not.

6.3.6.3 Scenario(s)

Referring to the use case described in clause 6.3.1, it is possible that personal data is available in both plaintext and ciphertext form. Namely, the freshest data resides in the device's secure execution environment, whereas the older data has been exported in ciphertext form to untrusted storage. Hence, a subject requesting data can either obtain plaintext data or ciphertext data. This is defined as direct access and indirect access, respectively.

Direct access is defined as the transfer of plaintext data from the device to the consumer. When the consumer connects to the device and requests data, the device determines that the requested data resides in its secure execution environment. The decision procedure to grant access to the data follows the usual approach of software-based access control systems where the consumer is authenticated, one or more access policies are selected, and the consumer's request is evaluated against the policies. The confidentiality of the data during transfer from the device to the subject is ensured using state-of-the-art communication security techniques.

Indirect access involves one level of indirection as the consumer will receive ciphertext. Instead of routing the request to software-based access control, the device determines that the data is old enough that it has been encrypted. Thus, it responds with the ciphertext corresponding to the request. The data consumer then proves that they have been granted an appropriate access policy in order to obtain the data in plaintext, by providing an appropriate key to the decryption step.

It is possible that the consumer receives a response involving direct and indirect access to data, depending on the request.

Two variants of the above scenario are possible:

- in the offline case described in clause 6.3.5; or
- in an online case, a network infrastructure (not necessarily the Internet) is available and reachable by devices, and there exists an entity in the infrastructure that is able to locate encrypted data on untrusted storage, as well as devices that can provide the most up-to-date data. When said entity receives a request for data, it gathers relevant ciphertext and, if necessary, forwards the request to relevant devices - or redirects the subject to relevant devices.

Example of the offline case is that of a contractor monitoring the activity log of a device as part of a maintenance visit.

Examples of the online case include remote monitoring of equipment as well as real-time retrieval of a device location in case of an emergency.

6.3.6.4 Information Flows

Pre-conditions: a cryptosystem for access control is configured for a set of devices, users and policies.

Post-conditions:

- Success: access control operates as expected and the data consumer receives the requested data as plaintext (direct access) or the data consumer receives the requested data as ciphertext (indirect access).
- Failure: the data consumer does not receive the requested data because of a system failure.

Begin triggers: the data consumer issues a data request.

Termination triggers: N/A.

Working assumptions: access control policies are pre-provisioned on devices. In an online case, a network infrastructure (not necessarily the Internet) is available and reachable by devices.

Normal flow: The direct access flow is described as follows:

- The consumer requests access to a dataset.
- The consumer is authenticated.
- One or more access policies are selected.
- The consumer's request is evaluated against the policies.
- The confidentiality of the data during transfer from the device to the subject is ensured using state-of-the-art communication security techniques.

Alternative flow 1: The indirect access flow is described as follows:

- The consumer requests access to a dataset.

- The device determines that the data has been encrypted.
- The device responds with the ciphertext corresponding to the request.
- The consumer proves that he or she has been granted an appropriate access policy in order to obtain the data in plaintext, by providing an appropriate key to the decryption step.

Alternative flow 2: The request does not match any known dataset:

- The consumer requests access to a dataset.
- A not found response is returned.

6.3.7 Access control examples in the Industrial Internet of Things

6.3.7.1 General use case description

The present use case provides several examples of data access control based on the characteristics of the data and of the data subject, in particular relative to: time, absolute position or relative position against one person and/or one machine, origin, data type, clearance level, subject role, and service tier.

6.3.7.2 Stakeholders

User: the user is a data subject.

Device manager: the device manager can coincide with the user or be a different stakeholder.

Data consumer: the consumer is a data processor.

Device: the device takes decisions on the ciphertext annotation based on the data characteristics and operating conditions.

6.3.7.3 Scenario(s)

In the scenarios below a basic configuration is considered where a device (such as a worker wearable or an industrial equipment) is able to associate operational data with a position and/or a timestamp. Alternatively, a third-party service provider is considered which can process operational data into generated data and can associate it with the position where the data was processed, and/or a timestamp. Additional metadata can be associated, depending on the access control scenario as described below. The example restrictions are not meant to be exclusive but rather complementary.

Examples of time-based access control:

- **Industrial wearable:** a worker is given a wearable device that will monitor environmental data for safety purpose (for example, the worker location, its health status, air composition or radioactivity levels). The wearable can generate data outside working hours. For example, it may be convenient to use and the employer can allow that the worker uses it outside working hours, for personal purposes. The employer only accesses data that has been captured during working hours. When the worker is a Lone Worker, a procedure exists so that this time restriction is waived when there is presumption of an accident.
- **Corporate car:** an employee is given a car by the employer for the conduction of its business (e.g. the employee is a sales representative or a worker that attends to various industrial sites). The employee can also use the car for personal purposes as a company benefit. As part of the work contract, the employee consents that the employer has access to data (such as location) that was captured during working hours (but not outside working hours). In contrast to this, the employee consents that the insurance company has access to the complete set of location data, for risk evaluation purpose.
- **Access by third-parties:** a service provider, such as an auditor, is granted access to a dataset that covers a limited period of time (e.g. the first quarter of a given year), so that they can perform the audit.

- **Generated Data:** a service provider performing data analytics has been granted access to Operational Data as well as the right to resell the Generated Data under certain conditions. In order to support their business model (subscription-based access), said service provider can associate a timestamp with the Generated Data and grant their customer access to data generated over specific time periods.

Examples of position-based access control:

- **Restricted areas:** a company owns engineering offices as well as factories. The company is active on the Consumer Equipment market but is also a contractor for the local military, defence, or civil security authorities. In another example, the company owns a site that is covered by safety regulations (e.g. EU Directive 2012/18/EU [i.22], also known as Seveso-III). Therefore, certain sites of the company are restricted areas, and the distribution of related operational and generated data is similarly subject to restrictions. Such data is associated with location metadata to help managing its distribution.

- **Industrial wearable:** referring to the industrial wearable example given above for time-based access control, it is acknowledged that the wearable can generate data at the workplace and out of the workplace. The operational data is associated with location metadata, so that the employer can only access the data that was captured at the workplace (identified e.g. as a set of shapes made of coordinates, or by the presence of a specific radio signal).

- **Different work regulations rules areas:** work regulations can be very different from one country to another one, as well as lawful interception rules. An employer can access data in one country easily, but the same data can be impossible to get without employee explicit consent in another country.

NOTE 1: While the position can be understood as a set of coordinates (e.g. longitude, latitude, and altitude) it can also be expressed in terms such as building, floor, room, or zone. The position can also be "absolute" when the reference is static or "relative" when there is no static reference.

NOTE 2: Time-based and position-based access control can be associated and together enable location-based access control.

Example of origin-based access control:

- **Data aggregation:** in a factory, a service provider is tasked with aggregating data from different equipment. The data is associated with its origin (a specific equipment) so that other actors can access it according to the need-to-know principle. For example, the service provider tasked with maintenance of a given category of equipment, can only access data that is originating from said pieces of equipment.

Examples of access control based on data type:

- **Aggregated data:** when data is saved to untrusted storage, it can be aggregated. However, not all actors can access all data. For example, access can be restricted to specific types (temperature, pressure, etc.), specific category of sensors (e.g. pressure sensors located in a specific zone) and other semantics that characterize the data itself. Data type can be further specialized, for example health data can be described as blood pressure, heartbeat, body temperature, and so forth). In a chemical plant, temperature data can be specialized into internal temperature and further to the temperature of the chemical reactor (the latter operating temperature potentially being an industrial secret).

- **Local operational data:** this case is similar to the previous one in the sense that a device can capture and generate different types of data (typically, a wearable can generate timestamps, location, proximity, and health data).

Mandatory access control:

- **Sites implementing a security clearance scheme:** referring to restricted areas described above, a company owning such restricted site can have equipment associate data (or a subset of the data) to a clearance level, such that an employee or other device wishing to access the data holds an access policy with the proper security clearance.

Role based access control:

- The previous case can be extended, whereby the data is also associated with a role identifier. This role identifier can be combined with a security clearance. For example, in a team of electrotechnical technicians, only a subset of the team can access data of power distribution panels connected to critical equipment. In another example, health data can only be accessed by life monitoring systems, company medical staff, and public medical staff dealing with the employee's health.

Access control based on the data subject:

- The previous case can be further extended so that the data is associated with its intended audience, e.g. an employee of a given branch/service, of a given rank and responsibility, or a software service of a given type.

NOTE 3: Recall that proper protection is necessary for attributes and policies which can disclose PII.

Access control based on service tier:

- **Generated Data:** a service provider performing data analytics can elect to associate Generated Data with a service tier in order to support their business model, so that different levels of subscription can lead to different sets of data and can be subject to different subscription fees.

NOTE 4: This type of categorization limits the distribution of data beyond the original purpose, and is only provided as an example for completeness.

Access control based on emergency level:

- The system has a notion of emergency level so that dangerous situations can be identified by devices (e.g. life-threatening situations, imminence of an industrial/environmental disaster, fire, flood, riots, etc.). In such case, access control on specific personal information, such as the data subject's location, can be waived. For example, during an industrial incident, first responders may need to access the worker's location data as quickly as possible, i.e. without any legwork.

6.3.7.4 Information Flows

Pre-conditions: a cryptosystem for access control is configured for a set of devices, users and policies.

Post-conditions:

- Success: access control operates based on one or more characteristics of the data.
- Failure: some characteristic of the data is not supported as an access control criterion.

Begin triggers: the data is being captured.

Termination triggers: the data has been encrypted.

Working assumptions: the device is able to determine relevant data characteristic for annotation.

Normal flow:

- Data is captured on the device.
- Prior to encryption, data characteristics (e.g. measurement timestamp or location, target roles, clearance levels, data type, and so forth) are determined by the device.
- During encryption, the data is annotated in the ciphertext based on the identified characteristics.
- At some point, data consumers are provided with cryptographic keys holding access control policies allowing access to the data according to their access rights.

Emergency situations, alternative #1:

- Data is captured by the device.
- The device determines that the data capture is done in while operating in emergency condition.

- The device identifies and applies the annotations relevant to the data characteristics as in the normal flow, along with an additional annotation indicating the emergency.
- At some point, data consumers are provided with cryptographic keys holding access control policies allowing access to the data according to their access rights in normal conditions.
- At some point, data consumers that have special access rights in emergency situations are provided with cryptographic keys granting access rights to data annotated with an indication of an emergency.

Emergency situations, alternative #2:

- Data is captured by the device.
- The device identifies and applies the annotations relevant to the data characteristics as in the normal flow.
- At some point, data consumers are provided with cryptographic keys holding access control policies allowing access to the data according to their access rights in normal conditions.
- During an emergency condition, data consumers that have special access rights in such situation request cryptographic keys granting access rights to data regardless of the annotation (or with less restrictive conditions as normal).

6.3.8 Recommended ABE schema

Depending on the required properties, either CCA-secure KP-ABE with KP-FAME-KEM (faster decryption) ABE scheme or KP-GPSW-KEM (slower decryption but allowing more complex policies and attribute repetition) ABE scheme [i.21] should be used.

7 Cloud use case

7.1 Introduction

7.1.1 Scenario

On the basis of indications provided by the European Digital Agenda, Italy has defined its own national strategy drawn up together with the Ministries and in collaboration with the Conference of Regions and Autonomous Provinces.

The Italian national strategy for ICT [i.1] leverages on the development of a number of enabling platforms, guidelines and development toolkits helping developers, designers and providers of third party services to build up their own services interoperable with the digital assets of the Italian public administration. Enabling platforms are solution that offers fundamental functionalities across different projects, simplifying development of services and reducing their costs. The development of citizen-centric services, in particular, will be based on a "mobile first" and personalized user experience.

Among different already available enabling platforms, the ANPR (Anagrafe Nazionale della Popolazione Residente) aims at unifying the many different demographic data management systems in use at the different Italian municipalities. Municipalities, other Public Administrations and citizens benefit from ANPR thanks to the de-duplication of entries (formerly duplicated on different databases) and to the real-time management of their data - as opposite to procedures involving citizens to provide different forms of attestations when moving from a municipality to another.

Clearly the ANPR contains both PII and personal data (citizens' family name, age, profession, address, family composition information, etc.) that need to be protected from unauthorized access. To this purpose, ANPR uses a Service Oriented Architecture relying on a PKI for providing authenticated access to client applications, SAML for expressing assertions and SOAP for message transportation.

To prevent access from unauthorized entities, ANPR embodies a Certification Authority (CA), providing certificates for the identity of each municipality, of each client application and of each device where the client application runs.

ANPR plays the role of Service Provider (SP) running a database containing citizens' data and providing web-based API to build client applications. Each operator using a device running a client application in a municipality premise plays the role of Service Requestor (SR).

While this security model permits to unambiguously identify and track each access (municipality, application and even single device basis), it may result difficult to scale up when thinking at ANPR as a possible core of a more general enabling platform providing services to a higher number of stakeholders. The latter can include the citizens themselves, which can encourage the development of third party applications providing electronic demographic services directly on their smartphones. As well, Government institutions and public service providers (e.g. public transportation, electric power providers, etc.) can benefit from accessing various kind of demographic information in order to provide citizens with value added services, to manage their infrastructures, to optimize local tolls, and to plan for further development of their infrastructure. These stakeholders can be interested in acting as citizen's data controllers, by obtaining part of the information contained in ANPR's database entries in various forms, in relation to their business and their goal (e.g. for their service execution, planning, statistical purposes, marketing research, etc.).

Clearly, there is a need to protect citizens' PII and their personal data, when these are transferred to third party providers. In order to comply with widely accepted privacy principles [i.20], as well as with national and European regulation [i.19], ANPR might set up different policy rules and efficiently give differentiated, finer granule permissions for client applications to access citizens' data based on their category, purpose and intended goals.

NOTE: This scenario has been selected as particularly fitting the definition of Platform Provider model. The "Platform Provider" model [i.21] is a specialization of the Identity Provider model described in ETSI TR 187 010 [i.16] offering additional data access functions, through API, to Relaying Parties. The "Device Platform Provider" model described in ISO/IEC 19944 [i.5] is a specialization of the "Platform Provider" model characterized by the presence of a (physical) specific device used as a client to access the cloud services.

7.1.2 Preliminary considerations

A recent trend for SOA services is to evolve and become part of an ecosystem, i.e. an architecture where a Platform provider offers identity services and API to access its services to a high number of loosely coupled third party stakeholders. These stakeholders act as service partners, building their own applications based on the API provided by the platform. Notwithstanding data portability, which is a vital feature in an ecosystem, the respect for citizens' privacy needs to be always ensured.

To satisfy various data access control needs, there may be the need to set up different policy rules and give differentiated permissions to third party's client applications based on their category and intended purposes, other than on their identity. The introduction of ABE can ensure an uniform method for controlling disclosure of data to third party client applications in a differentiated way, based not only on the identity of the service requestors but on various policy rules reflecting a finer-granule check of their purposes. The solution scales up with the number of applications as the Platform provider, using attributes, might group and classify each application in relation to its nature and objectives, and define policies to disclose data accordingly. For instance, ANPR might maintain identity certificates for each third party organization willing to access its database, but been relieved from maintaining certificates and permissions for each third party application and each device where the third party application runs.

In another example, services from different third party providers can be grouped based on the category of citizens' data they access and the goals of their processing. In this example, the "data use statements" format defined in ISO/IEC 19944 [i.5] can be used to express which data category from the ANPR the third party service uses and for what purpose. One such a statement might be: "A local bus company may use demographic information from ANPR to personalize bus fares". This statement would be translated into one (or more) assertions as defined in ETSI TS 103 532 [i.21], in particular:

- data access assertions, used in the evaluation of access control policies when the policy relies on properties of the subject (e.g. to describe the kind of client applications, its provider, its location, the purpose of the processing, actions carried out on data, etc.);
- data capture assertions, used during the evaluation of an access control policy when the policy relies on properties of the data (e.g. to characterize the kind of data, degree of linkability to an individual, etc.).

The ANPR then can build a policy granting access to encrypted data based on this statement and release secret keys corresponding to attributes fitting the involved properties.

7.2 High level requirements

The present document specifies the following high level requirements (table 7.1) for the Cloud use case.

Table 7.1: Requirements derived from the "Cloud" use case

ID	Requirement
CL#1	To protect the confidentiality of data, the solution shall use a cryptographic scheme supporting attributes, discouraging disclosure to unauthorized entities.
CL#2	The solution shall be compatible with the legacy infrastructure and shall comply with the standard roles defined in ISO/IEC 17789 [1].
CL#3	The solution shall encourage the development of an ecosystem, based on uniform API provided by the Platform provider. The solution shall scale with the number of third party applications which might be possibly unknown to the Platform Provider, still preserving a mechanism to provide access control over the data provided to the third parties.
CL#4	The provision of data to third parties may be based, other than on the identity of the third party service requesting the data, on various policy rules reflecting a finer-granule check of their intended purposes, complying with privacy principles of ISO/IEC 29100 [i.20]. The solution may support "data use statements" format defined in ISO/IEC 19944 [i.5].
CL#5	The solution shall allow for the provision of new secret keys on a regular basis, and shall allow for the verification of the recipient's properties.
CL#6	The solution shall support attribute revocation, in which case new attributes and secret keys shall be used. Revoked attributes shall not be used.
CL#7	The solution shall use a single, global, master public key MPK for encryption and a multiplicity of secret keys provided to authorized third parties for decryption.

7.3 Use case

7.3.1 Stakeholders

A Service Oriented Architecture running on a standard Cloud infrastructure [1] is assumed.

Cloud Platform Provider (PP): is a Cloud Service Provider providing identity management services and interfaces for third party applications using the platform services.

Cloud Service Partner (CSPa): provides support to the provisioning of cloud services by the Cloud Service Provider, or to the consumption of cloud service by the Cloud Service Customer.

Trusted Authority (TA): it has the following responsibility:

- generates the ABE master public key PK and the corresponding master secret key according to a selected large universe ABE scheme;
- supports the distribution of attributes used in encryption;
- supports the distribution of secret keys.

NOTE 1: There is a trust relationship between the PP and the TA (they can even be the same actor, as in the ANPR use case) and between the CSPa and the TA.

Cloud Platform user (Pu): is Cloud Service user consuming one or more platform services.

NOTE 2: There is a trust relationship between the Pu and the PP. The PP acts as an identity provider for the Pu. In addition, the PP acts as a PII processor for the Pu.

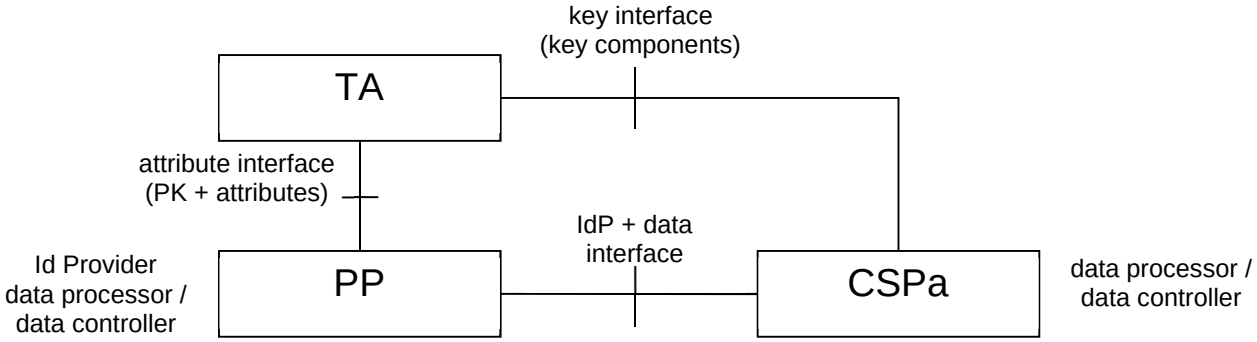

Figure 7.1: Architecture

7.3.2 Preconditions

The PP and the CSPa act as PII processor and provides processing of PII. In addition, the PP and the CSPa can act as PII controller, determining the purposes and means for processing PII.

The PP has received the ABE master public key (and ABE attributes) from the TA.

The CSPa is identified by a CSPa identifier. It has obtained a certificated identity from the PP.

The CSPa runs client applications (applications exploiting functionalities exposed by the platform, e.g. through an API).

The CSPa has obtained from the TA one or more ABE attribute(s) and related keys fitting the nature of its applications and purposes.

The CSPa authenticates its client applications.

7.3.4 Trigger

A client application issues a data request to the PP.

7.3.5 Flow of events

The CSPa generates an assertion (e.g. a SAML assertion) for the client application. The assertion contains: the CSPa identifier, the public ABE attributes assigned to the client application, and the CSPa certificate. By this request, the CSPa can formally assert data use statements.

The CSPa signs the assertion.

The assertion is inserted in the data request (e.g. a SOAP request header).

The CSPa signs the data request.

The data request is sent to the PP.

The PP checks the validity of the request, including the validity of included data use statements, if any. If the request is not valid a failure message is returned.

If the request is valid, the PP executes a query on its database to retrieve the requested data.

The PP selects a policy based on the received attributes and encrypts the result of the query using an ABE scheme. If the request contains a data use statement, the PP encrypts data using one or more policy fitting the data use statements.

The PP returns the ciphertext to the client application.

7.3.6 Exit condition

The client application decrypts the ciphertext using its ABE keys.

7.3.7 Recommended ABE scheme

CCA-secure CP-ABE with CP-WATERS-KEM and/or CCA-secure KP-ABE with KP-GPSW-KEM [i.21] ABE schemes should be used.

Annex A (informative): Attribute Based Encryption

A.1 Early ABE constructions

Traditional public-key encryption uses a public key to target ciphertexts to a specific user. The user holds a secret key and can decrypt the message. In attribute based encryption, instead, ciphertexts are not necessarily encrypted to one particular user. Instead both users' secret keys and ciphertexts can be associated with a set of attributes or a policy over attributes. A user is able to decrypt a ciphertext if there is a "match" between his secret key and the ciphertext.

NOTE: This feature enables a party to encrypt data even without knowing a priori the identity of each individual user which needs to access the data. ABE can be efficiently used whenever ciphertexts have to be targeted to a class of users rather than a single user, because encryption is performed once.

In a very early formulation of ABE [i.12], ciphertexts are labelled with a set of attributes S. A user's secret key is associated with both a threshold parameter k and its own set of attributes S′. In order for a user to decrypt a ciphertext at least k attributes overlap between the ciphertext and his secret keys [i.14].

A.2 Key Policy Attribute Based Encryption (KP-ABE)

Key Policy ABE [i.13] leverages on a more general attribute based encryption method, where a ciphertext is associated with a set of attributes and a user's key can be associated with a "monotonic tree access structure". In a monotonic tree access structure, each non-leaf node of the tree is called a "threshold gate". Each leaf is associated with attributes.

In an access structure, Boolean operators "AND" and "OR" can be represented through gates using appropriate threshold. In particular, a Boolean "AND" operator is represented by a gate with a threshold set to the number of its children. A Boolean "OR" operator is represented by a gate (with more than one child) with a threshold set to 1.

A set of attributes is said to satisfy an access tree if:

- the access tree consists of only one node, labelled with an attribute, and the set of attribute contains that attribute; or
- recursively, each subtrees rooted in the root of the tree is evaluated until its leaves. Then, check whether the threshold is matched (in case of OR at least one of the subtree is satisfied, in case of an AND all of its subtree are satisfied).

In KP-ABE a monotonic tree access structure representing an access policy is encoded into the users' secret keys. The ciphertext is computed with respect to a set of descriptive attributes S. A user can decrypt the ciphertext if and only if his/her secret key encodes the correct access structure. The secret key's access structure therefore specifies which ciphertexts the key holder is allowed to decrypt.

In KP-keypolicy ABE, the encryptor exerts no control over who has access to the data it encrypts, except by its choice of descriptive attributes for the data. Rather, it trusts that the key-issuer issues the appropriate keys to grant or deny access to the appropriate users. The "intelligence" is assumed to be with the key issuer, and not the encryptor [i.14].

NOTE: This formulation makes KP-ABE appealing whenever access to a given resource has to be granted based on attributes the resources is labelled with.

The whole cryptosystem consists of four algorithms (functions):

- Setup(l, U): Takes the global parameter l and an attribute universe U description as input and outputs the public parameters MPK and a master key MSK.
- Key-Gen(MSK, A): Takes as input the master secret key MSK and an access structure A over a set of attributes S that describe the key and outputs a secret key SK.
- Encrypt(MPK, M, S): Takes as input a message M, a set of attributes S, and the public key MPK, and outputs the ciphertext CT.

- Decrypt(CT, SK): Takes as input the ciphertext CT that is encrypted under the set of attributes S; the secret key SK, that represents the access policy A. It outputs the original message M if and only if S satisfies the access policy A.

A.3 Ciphertext Policy Attribute Based Encryption (CP-ABE)

CP-ABE is dual to KP-ABE, meaning that a monotonic tree access structure representing an access policy is encoded into the cyphertext, while the user's secret keys are computed with respect to a set of attributes S' the user has obtained.

In CP-ABE, attributes are assigned to users and the access structures are used to label different sets of encrypted data. A user is able to decrypt the ciphertext with a given key if and only if there is an assignment of attributes from the secret key to nodes of the tree such that the tree is satisfied [i.14].

The underlying mathematical construction is based on a secret sharing scheme embedded in the ciphertext that prevents collusion attacks, i.e. attacks from two or more different users that have obtained keys encoding different sets of attributes that, by their own, would not satisfy the access tree structure, but whose union would do.

NOTE: A dual property holds for KP-ABE, where a secret sharing scheme is embedded in the secret key.

The decryption algorithm works by masquerading original values from each particular user's secret key in a randomized fashion. Thus the key components associated to the same attribute in two different secret keys are not the same and cannot be interchanged in order to decrypt a cyphertext.

The scheme consists of four algorithms:

- Setup(l, U): Takes global parameters and attribute universe U description as input and outputs the public parameters MPK and a master key MSK.
- Key-Gen(MSK, S): Takes as input the master secret key MSK and a set of attributes S that describe the key and outputs a secret key SK.
- Encrypt(MPK, M, A): Takes as input the public parameters MPK, a message M, and an access structure A over the universe of attributes (that represents the policy to be satisfied to access the message). The algorithm encrypts M and produces a ciphertext CT such that only a user that possesses a set of attributes that satisfies the access structure will be able to decrypt the message. The ciphertext implicitly contains A.
- Decrypt(CT, SK): Takes as input the ciphertext CT, which contains the access structure A, and a secret key SK, which is a secret key for a set S of attributes. If the set S of attributes satisfies the access structure A then the algorithm will decrypt the ciphertext and return a message M.

A.4 Key distribution protocols

A method to use a traditional Public Key Infrastructure for distributing ABE keys inside X.509 certificates and maintaining Certificate Revocation Lists is described in [i.21].

A.5 Attribute revocation

Attribute revocation consists in not allowing a user to use old decryption keys associated with lost or revoked attributes. A trivial way of implementing revocation would be to change all the decryption keys and re-encrypt all data items. However, this solution is not practical. In the literature, two main categories of solutions have been proposed:

- attaching expiration dates to the attributes, in such a way that decryption keys cannot be used for encryptions produced after the expiration date; or
- employing a proxy that generates a decryption token after verifying that there are no revoked attributes.

The first approach has the drawback that the decryption key cannot be immediately revoked when the user loses some of his attributes, but it will be expired after passing certain periods.

The disadvantage of the second approach is that the proxy needs to be fully trusted so that a token for a user with revoked attributes is not generated; an additional drawback is that the proxy needs to perform heavy computations scaling with the number of access requests.

A.6 Key expiration approach

The key-expiration approach uses a time based decryption key, so that the key expires after a certain period of time. Another approach [i.17] consists in attaching an expiration date to the attributes of the decryption key so that this key cannot be used for decryption after the attached date expires. When a key reaches the expiration date, the key generation authority checks its revocation list and updates only the decryption keys of users who do not have any revoked attributes. However, this approach requires that either the expiration date of decryption keys is the same for all the issued keys in order to enable the key authority to update the keys only at specific times, or the key generation authority updates the key frequently. This requirement makes the system either less efficient or less flexible.

To tackle this problem [i.15] proposed a scheme that allows the key authority to update the decryption keys more efficiently. However, this approach is proposed for Identity Based Encryption schemes, a particular case of ABE where the attribute is the identity of the recipient.

Bethencourt [i.14] proposed a solution that attaches expiration dates to every attribute of the decryption key and achieves enforcement by doing numerical comparisons in the policy tree. However, this scheme increases the complexity of creating the ciphertexts and the decryption key because each attribute is replaced by n attributes, where n is the number of bits used to represent the expiration date. In practice to implement a numerical comparison using policy trees one needs n more leaves (attributes) for an n-bit number. This method requires also that the user encrypting the data synchronizes the encryption time with the key generation authority so that both the encryption and decryption use the same time.

There are two major problems with the key-expiration and periodic key updates approach. Firstly, the attributes can only be used for new encryptions, because for older encryptions a user can still use his old decryption key. Moreover, revocation only goes into effect for encryptions produced after the new expiry date, which is usually not immediate. The complexity of the policy tree also grows unjustifiably, which can make the encryption and decryption infeasible.

A.7 Mediator approach

The mediator approach is based on proxy re-encryption methods which use a proxy to re-encrypt the ciphertext without revealing the plaintext. This method involves a third semi-trusted party that performs computations under the honest-but-curious trust model.

Ibraimi et. al [i.11] propose a proxy mediator approach where the decryption key is broken into two parts: the proxy receives one part of the decryption key and the user receives the other part. To decrypt data, the user has to first send the ciphertext to the proxy which generates a decryption token using the proxy key. The proxy holds an attribute revocation list which is updated as soon as a user loses some of his attributes. The proxy generates the token only if user attributes do not occur in the revocation list. Revocation is therefore enforced using the revocation list maintained by the proxy. The main advantage of the proxy approach is that when an attribute is dropped from a user, the attribute will be immediately revoked form the decryption key of the user. However, this approach has inefficiency drawbacks. The proxy mediator is required to decrypt data partially upon receiving each request to generate a token. These computations impose a high load of computations which can make the proxy a bottleneck in the system when a large number of users are involved.

User and attribute revocation can be achieved with proxy re-encryption using a delegation key [i.18]. The delegation key has two functions:

i) re-encrypt every ciphertext requested by legitimate users; and

ii) can be updated to reflect the user attributes revocations.

In cyphertext splitting [i.6], the ciphertext of the ABE scheme is divided into two parts, where one part is stored on the storage cloud and the other part is stored on a mediator proxy. The proxy is a semi-trusted entity that in addition to holding the corresponding part of the ciphertext has a revocation list that contains the last status of the user attributes. Each time a user wants to access the data, he needs to download the first part of the ciphertext from the storage cloud and then the second part from the proxy. Before sending the second part of the ciphertext, the proxy checks whether the user has any revoked attribute. If this is the case, the proxy does not send the second part of the ciphertext to the user, which prevents decryption of the data.

The ciphertext splitting using a proxy solution allows immediate revocation of attributes. Furthermore, it is more computation-efficient than other existing proxy mediator systems. On the other hand, in contrast to the previous proxy mediator systems, in this solution the proxy needs to store a part of each ciphertext instead of just key material, increasing its storage costs.

A.8 Relationship with Attribute Based Access Control (ABAC)

As ABE provides implicit authorization, eliminating the need for a software system implementing evaluation of policy and decision enforcement to access to data, the term "encrypted access control" is sometimes used to refer to it. Leveraging on pure mathematical algorithms, encrypted access control can protect the confidentiality of data even if the software infrastructure is compromised. This feature makes ABE appealing for access control enforcement in distributed systems where limited trusted software, firmware or hardware exists, or when it is impossible or impractical to adopt an online access control system. Nevertheless there are limitations that need to be considered when thinking at encrypted access control as an alternative to traditional access control mechanisms. [i.21] presents a list of well-known limitations.

Annex B (informative): Compliance with Lawful Interception principles

Whenever applicable (i.e. whenever a national or international regulation exists), the present document preserve two fundamental principles from Lawful Interception [i.10]:

- LI#1: The Law Enforcement Authority (LEA) should be able to identify and communicate with the service provider(s) responsible for the communications involving specific targets.

NOTE 1: In the present document this principle is respected because the entities providing services (mobile network operators, cloud service providers, etc.) are registered entities subject to national and international regulation.

- LI#2: The service providers who initiate encryption should be able to provide intercepted telecommunications en clair or provide the LEA with the keys and other information needed to access the information.

NOTE 2: In the present document such keys and information are always available from both:

1) the trusted authority (all released keys); and

2) the service provider responsible for decryption (limited to the keys received from the trusted authority, which are used for its own business).

NOTE 3: The present document does not give end-user control of end-to-end encryption, rather whilst ABE enables end-to-end encryption the keys are always known to the trusted authority, acting as a key escrow entity, such that the service provider is able to ensure any authorized and necessary access to communication for lawful purposes.

History

Document history		
V1.1.1	June 2018	Publication