Specific Targeted Research Projects (STReP)

SOCIOTAL

Creating a socially aware citizen-centric Internet of Things

FP7 Contract Number: 609112

WP3 – Privacy-aware communication

Deliverable report

Contractual date of delivery:
M18 - 28/02/15
Actual submission date:

Deliverable ID:	**D3.1.2**
Deliverable Title:	**Device centric enablers for privacy and trust**
Responsible beneficiary:	CEA
Contributing beneficiaries:	UNIS, CEA, UMU
Estimated Indicative Person Months:	8

Start Date of the Project: 1 September 2013 Duration: 36 Months
Revision:
Dissemination Level: Public

Document Information

Document ID: 3.1.2
Version: V1.6 (Final)
Version Date: 05 March 2015
Authors: **Michele Nati, Niklas Palaghias (UNIS), Benoît Denis, Iulia Tunaru (CEA), Victoria Moreno Cano, Jose Luis Hernandez Ramos, Jorge Bernal Bernabé (UMU)**
Security: **Confidential**

Approvals

	Name	Organization	Date	Visa
Project Management Team	Klaus Moessner	UNIS		

Document history

Revision	Date	Modification	Authors
0.1	30/10/14	First ToC	Benoît Denis
0.2	28/11/14	Revised ToC after Michele's comments and 1srst review feedback	Benoît Denis
0.3	19/12/14	Draft section 9 for enablers security assessment (examples)	Benoît Denis
0.4	09/01/15	Revised sections 2.2 (SotA) & 9.2 (Security) related to radiolocation enablers	Benoît Denis
0.5	14/01/15	Integration of first inputs from UMU and UNIS (merged version)	Victoria Moreno
0.6	27/01/15	Updated sections 3.2 (scenarios), 8.2 (WP2/WP3 integration) related to radiolocation enablers	Benoît Denis
0.7	02/02/15	Updated sections 5 (enabler description) & 7.2 (Evaluation)	Benoît Denis
0.8	05/02/15	Updated sections 3.1 (scenarios), sections 4 (enabler description) & 8.1 (WP2/WP3 integration) related to F2F enabler	Michele Nati, Niklas Palaghias
0.9	05/02/15	Merged version with ToC refinements and new responsibilities assignment after WP2/WP3 PhoneCall	Benoît Denis
1.0	16/02/15	Updated sections 3.2 (scenarios) & 7.2 (Evaluation) regarding radiolocation enablers; Added section 9.4 (Overall security assessment)	Benoît Denis
1.1	17/02/15	Added section 1 (Main introduction)	Benoît Denis
1.2	20/02/15	Added sections 7.4 (Overall evaluation), 8.1 (WP2/WP3 Integration) related to F2F enabler & 10 (Overall conclusion).	Michele Nati

1.2_UMU	20/02/15	Revised sections 7.3 (Evaluation) & 8.3 (WP2/WP3 integration) related to magnetic-based indoor localization	Jorge Bernal Bernabé
1.2_UNIS	22/02/15	Overall revision of sections related to F2F enabler (based on inline comments)	Niklas Palaghias
1.3	23/02/15	Revised section 8.2.4 (WP2/WP3 integration) related to radiolocation enabler, added executive summary	Benoît Denis
1.4	27/02/15	Document review	Dejan Drajic
1.5	03/03/15	Document review	Sutharshan Rajasegarar
1.6 - Final	05/03/15	Revised/consolidated version after reviews	Benoît Denis

Content

Figures

Tables

Acronyms and Abbreviations

AP	Access Point
BF	Brute Force
BS	Base Station
CapBAC	Capabilities Based Access Control
CoAP	Constrained Application Protocol
CP-ABE	Ciphertext-Policy Attribute-Based Encryption
CSI	Channel State Information
DHC	DARSIS Hierarchical Classifier
DOE	Dynamic Offset Estimation
DPC	DARSIS Proximity Classifier
DSC	DARSIS Single Classifier
F2F	Face-to-Face
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSM	Global System for Mobile communications
ID	IDentity number
IoT	Internet of Things
IR-UWB	Impulse Radio - Ultra Wideband
LDR	Low Data Rate
LBP	Location - Based Pseudonym
LBK	Location - Based Key
LoS	Line of Sight
LQI	Link Quality Indicator
LT	Location and Tracking
MAC	Medium Access Control
NB	Narrow Band
NLoS	Non Line of Sight
NLoS2	severe Non Line of Sight
NN	Nearest Neighbour
pBF	probabilistic Brute Force
PCA	Principal Components Analysis
PLM	Path Loss Model
PUF	Physically Unclonable Functions
P2P	Peer-to-Peer
RBF	Radial Basis Functions
RFID	Radio Frequency IDentification
ROC	Receiver Operating Characteristic
RSSI	Received Signal Strength Indicator
RT-ToF	Round Trip - Time of Flight
SA	Similarity based Authentication
TDMA	Time Division Multiple Access
TDoA	Time Difference of Arrival
ToA	Time of Arrival
T&R	Trust & Reputation
T&RM	Trust & Reputation Manager
TRVA	Temporal RSSI Variation Authentication
TWR	Two Way Ranging
VE	Virtual Entity
WSN	Wireless Sensor Network

Executive summary

This deliverable describes the final efforts and achievements of SocIoTal towards the definition of new device-centric enablers for privacy and trust. Three main enablers have been proposed, along with various embodiments and variants. The first one, namely the Face-to-Face (F2F) interaction enabler, allows to detect and identify F2F social contacts among users by relying on the observed interactions between their mobile (smart) phones. A second enabler leverages radiolocation techniques in order to provide devices with new identity mechanisms. This enabler, which aims at preventing impersonation attacks through location-based pseudonyms (LBP) verification, is primarily intended as a dynamic protection overlay to assist conventional authentication procedures. Finally, a third enabler can accurately characterize devices and users positions in indoor environments, by sensing geo-magnetic field variations at smartphones. The retrieved location information is accessible through a specific service and directly exploitable as part of the context data, similarly to the radiolocation enabler. But it can be also used to further grant authorization to the users or devices based on their physical locations, for instance in resource access control.

The three proposed enablers comply with the heterogeneity and capabilities of the devices envisioned in SocIoTal scenarios, considering typically (but not restricting to) end users with their associated smartphones. The later devices can for instance be used as gateways to access and contribute to services by generating and sharing relevant data, but alternatively they could be a variety of embedded systems, either provided by the users (still) or by other entities such as existing Smart Cities elements of infrastructure. In fact, while possible adaptation of the F2F or magnetic field sensing enablers could be envisioned with other kinds of devices (rather than smartphones), likewise the adaptation of radiolocation enablers to different radio technologies is fully supported.

The final definition, description, evaluation and integration of the three proposed device-centric enablers address the Objective O3.1 defined for Work Package 3 and reported as **O3.1: To develop novel device-centric enablers for the automatic establishment and enforcement of identities and trust relationships**. These investigations first concern techniques for the extraction of device-level secure and hard to tamper identity profiles, for instance based on the physical location. New solutions are also put forward for the detection of face-to-face interaction patterns between humans based on mobile devices. The later can thus be used as foundation to further infer trust relationships. Therefore, with the definition of the three enablers, it can be concluded that the initial aims of the Task T3.1 have been met.

The previous enablers are herein presented according to a common and systematic structure, for better readability. Besides final improvements, in comparison with the intermediary version of this document (D3.1.1) [5], the focus is put on consolidated validations through novel experiments and simulations, on more concrete use cases, on the integration flow with respect to other SocIoTal components and on intrinsic enabler security.

The deliverable is organized as follows. First, each single contribution is positioned with respect to recent state-of-the-art, recalling not only their expected benefits separately, but more generally, justifying the retained device-specific approach in comparison with more conventional solutions. Then, high-level application scenarios are exemplified into inheriting use cases, illustrating enablers' practicability and concrete utility, as well as application-specific challenges. The latest enabler additions and improvements are subsequently detailed, from both architectural and implementation perspectives. After that, the upgraded enablers are evaluated based on both field experiments and extensive scenarios in realistic operating contexts. Then, the interfacing and integration of each enabler -as a "tool"- with respect to *trust and reputation* or *secure communication* components is

shown for the specified application scenarios, in compliance with the SocIoTal context model and overall architecture framework. On this occasion, the way the information provided by each enabler is exposed to other architecture components is detailed, as well as the way such information can be exploited in order to support the functionalities envisioned by the SocIoTal platform (e.g., so as to infer new trust relations, provide new context definitions, complete new identity profiles, support new access control…). Finally, the security of each enabler is first assessed independently through a threat analysis, before pointing out reciprocal synergies between enablers, as well as their key role into the holistic SocIoTal security paradigm.

The three main enablers, considered as independent building blocks, can be summarized as follows.

First of all, the technique for *Face-to-face interaction detection* uses mobile smartphones' interactions to detect when two people are engaging in a face-to-face interaction by combining their interpersonal distance and their facing direction, independently of external hardware and internet connectivity conditions. This is achieved by applying techniques that solve out the *Direction detection* problem regardless of the on-body position of the mobile device. Additionally, a technique for logging multiple Received Signal Strength Indication (RSSI) measurements from the same device is also provided in order to collect a larger amount of samples in a short time frame and to detect *Nearby devices*. The latest provided evaluations relying on representative simulations (concerning also the newly added inter-distance estimation feature) show interesting performances in comparison with state-of-the-art approaches.

The second enabler, which exploits radiolocation resources and modalities available at most IoT devices (e.g., Global Navigation Satellite System, Received Signal Strength Indicators or Round Trip - Time of Flight over short-range wireless communication links, etc.), allows pseudonyms generation and verification out of radiolocation data to prevent impersonation attacks. More precisely, techniques are proposed to make use of relative peer-to-peer ranging information (between neighbouring devices) and optionally, absolute devices' 2D positions, to locally produce link-specific and/or position-specific pseudonyms. The latter can be used in dynamic authentication procedures through pseudonyms verification (e.g., during discovery phases while trying to establish areas of trust), and respectively in ad hoc cases through legitimate pseudonym guess or relying on a centralized entity (e.g., IoT context broker). Evaluations based on indoor measurements issued at real radio devices have been considered together with simulations, confirming the accuracy and resilience to security attacks of the proposed solutions in real application scenarios.

Finally, the third proposed indoor localization enabler is based on the use of sensors, which are integrated in common smartphones. Therefore, unlike most of current proposals, this stand-alone approach does not require the deployment of additional hardware, devices or elements of infrastructure, thus providing a flexible and easy manageable indoor system and service, not only to end users (i.e., for indoor navigation purposes, as a service) but also to feed advanced location-based access control mechanisms. The enabler performance has been evaluated and compared in real-world environments against existing state of the art solutions as well, showing suitable accuracy into the envisioned SocIoTal application scenarios.

Overall, based on both initial and final evaluations, the designed enablers seem to show enough maturity and improvement to guarantee an adequate level of quality for the information they will generate and provide to other SocIoTal components. Their upcoming integration in the SocIoTal platform and in end-to-end field trials should also provide more tangible feedback on different evaluation aspects not fully covered so far, such as energy and computational efficiency, in order not to compromise system performance and user's experience.

Section 1 - Introduction

This deliverable accounts for the latest and final achievements of Task 3.1. The set of **device-centric enablers for privacy and trust** initially introduced in D3.1.1 [5] has been upgraded and evaluated through novel experiments, hence reaching more conclusive and realistic performance assessment (though still remaining at the enabler level), as well as more practical implementation. Specifically, the three proposed device-centric enablers concern Face-to-Face (F2F) interactions detection and handling, impersonations prevention through radiolocation-based pseudonyms (LBP) generation and verification and finally, indoor localization through magnetic field sensing at smartphone (for e.g., access/authorization control).

As a major outcome of this task, the later enablers are seen as key building blocks and mechanisms that enhance the identity and trust relationship management, which represents the core of the SocIoTal framework developed in WP2. Moreover, the same enablers, together with a set of rules conceptualizing the "trust zone" and "community" defined in WP2 (see also Deliverable D2.1 [66], support real-time context acquisition and identification at the device (and optionally, at the back-end level), as well as the identification and selection of appropriate communication behaviours in terms of discovery and routing. The inference of specific contexts, relying on the information produced by the considered enablers, will be implemented and provided as input to a context manager accessed by the **Privacy-aware communication framework** developed in T3.2. Figure 1 shows interactions between the three technical Work Packages of the project (WP2, WP3 and WP4), highlighting the central role of Task T3.1.

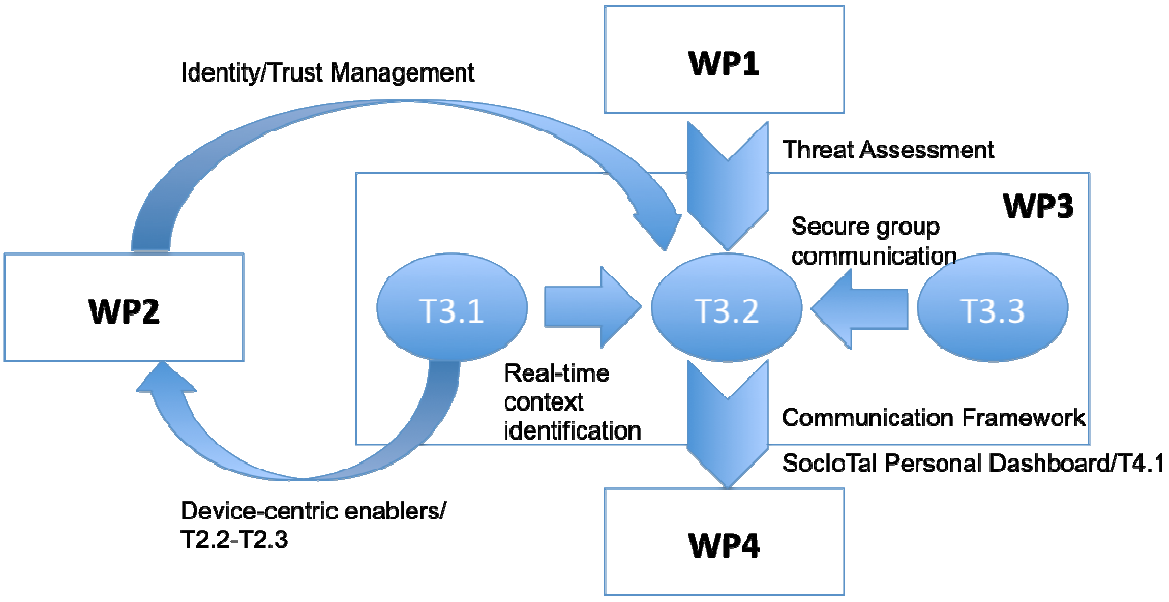

Figure 1. WP3 tasks dependencies

In comparison with D3.1.1 [5], a refined description of the enablers implementation and improvements is provided. Beyond, a global integration path is illustrated within the overall SocIoTal framework, showing interfaces and further usage with other management components (in compliance with the defined APIs and overall context model). On this occasion, practical high-level scenarios and use cases are also described. Finally, the security of the proposed enablers is assessed through a complete risk analysis, emphasizing the holistic approach of SocIoTal in terms of security

and privacy, but also synergies between the intrinsic enablers' properties and more conventional security approaches (e.g., cryptography, pseudonyms).

Section 2 - Motivations and progress beyond State-of-the-Art

The following section presents a review of current state of the art relevant to every considered enabler, focusing mostly on the progress realized beyond and justifying the choices made with respect to more conventional device-centric approaches.

2.1 Face-to-face enablers

In this section we present state-of-the-art solutions for detecting social interactions. Initially researchers tackled the problem of estimating on-going social interactions through custom devices that were based on Infrared [31], [32] and RFID [33] technologies. These devices achieved high accuracy but suffer from some important limitations including lack of pervasiveness, wearing position restriction, need of external hardware and do not allow large-scale deployment. To overcome these barriers we focus on off-the-shelf mobile phone solutions.

2.1.1 Coarse-grained distance estimation

The existing trend for inferring social interactions on mobile phones is detecting devices in close range. In [34], *CenceMe* [35], *Serendipity* [36] and *SoundSense* [37] authors discover nearby devices through Bluetooth, GSM signals or WiFi and classify them as a social interaction. However, these approaches induce a considerable amount of error, which is proportional to the range of the communication interface.

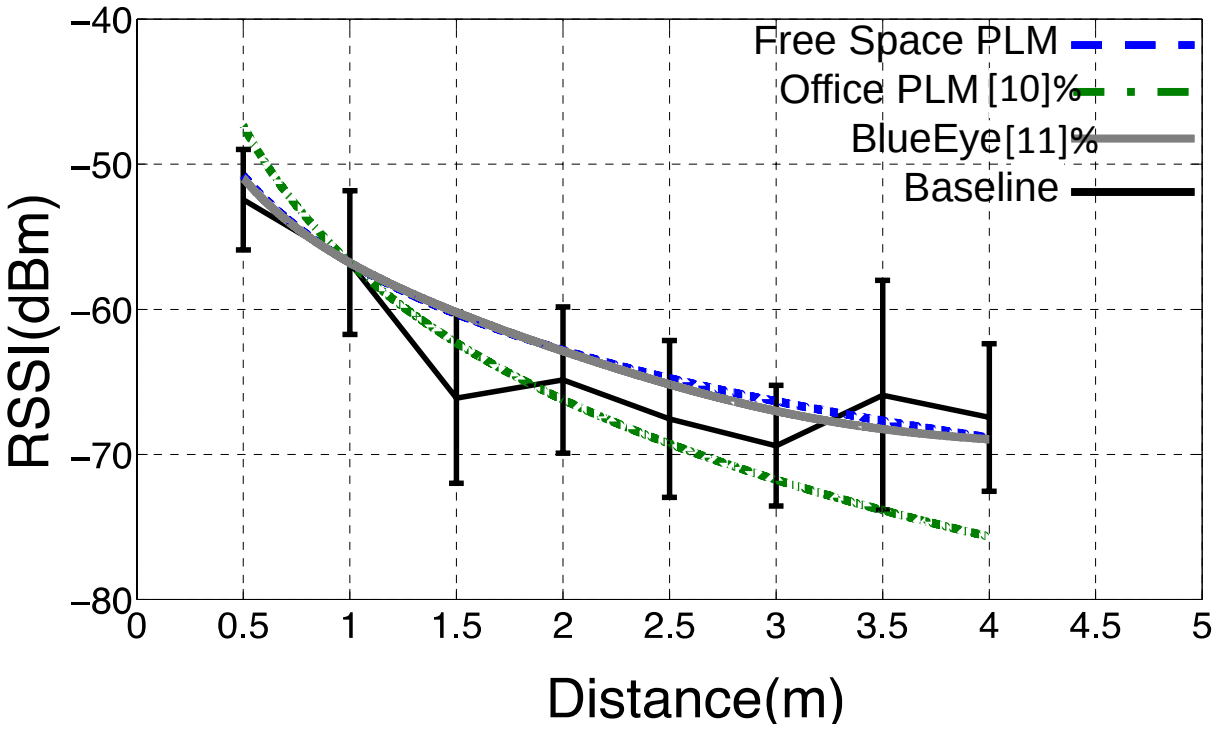

Figure 2. Comparison of different propagation models.

2.1.2 Fine-grained distance estimation

To improve the accuracy of detecting social interactions through simple device discovery, researchers focused on estimating interpersonal distance among users and classifying if they were in proximity or not. Several ways of estimating distance on smartphones have been proposed including Time of Arrival, Time Difference of Arrival, Angle of Arrival and RSSI. However, we focus on RSSI due to its low

complexity and pervasiveness as opposed to other approaches that require firmware modifications or external hardware. A significant part of proximity detection has focused on leveraging different Bluetooth RSSI path loss models (PLM) such as *Free Space PLM*, *Office PLM* [38] and *BlueEye* [39]. But as shown in Figure 2 an initial evaluation indicates low accuracy, which is justified by models' generality. *Comm2Sense* [40] performed proximity detection offline based on a 20-sample WiFi RSSI window fed in a naive Bayes model with kernel density estimator [65]. Mobile phones required firmware modification to switch on/off WiFi hotspot mode and set transmission power to 0dBm. *PhoneMonitor* [41] detected social interactions through a probabilistic model based on Bluetooth RSSI. *MAUC* [42] estimated through simple Bluetooth RSSI threshold if two people were in contact, by considering also user's movement. These approaches do not consider relative orientation (i.e., facing directions) of the users. Further, they perform their analysis offline, allowing only retroactive detection of social interactions and introducing various user privacy issues as the data are retrieved from the device and processed at an external server.

2.1.3 Multimodal inference

For developing more accurate methods researchers focused on incorporating multiple modalities. *Virtual Compass* [43] utilizes RSSI values (mean and uncertainty) from Bluetooth and WiFi interfaces to build regression models based on pre-acquired training set, for estimating distance and relative spatial arrangement among devices. Because they operate with both Bluetooth and WiFi interfaces in the inference process, energy consumption constitutes an important issue that prohibits the usage in real-world applications. Further, the relative orientation concept for detecting a social interaction cannot be inferred from relative spatial arrangement. An advancement of *Comm2Sense* [40] led to a recent attempt towards social interaction detection on mobile phones presented in [44]. The approach combines proximity detection, calculation of relative orientation and speech activity detection through external accelerometer attached to user's chest. However, it falls short in several aspects. Social interaction inference was executed offline while the device was placed at a predefined on-body position. A common drawback shared with *Virtual Compass* is that to this moment switching on/off WiFi hotspot mode requires firmware modifications. Although successful in their domain, existing techniques are facing several challenges limiting their applicability in real-world environments. In the following subsection we introduce an initial attempt to resolve these problems and provide an accurate and non-intrusive tool, suitable for long-term monitoring of social interactions.

2.1.4 Our approach

To resolve the shortcomings of previous approaches, we design, implement and evaluate a novel opportunistic approach for detecting social interactions by using only off-the-shelf mobile phones, which does not require user involvement in the inference process. Initially we develop a novel method for detecting users' interpersonal distance in a fine-grain manner based on only 6 Bluetooth RSSI samples. We estimate users' relative orientation by improving a state-of-the-art technique [30] for facing direction detection, independent to device on-body wearing position and enhance these through collaborative sensing for faster, privacy preserving and real-time inference. Additionally, the face-to-face enabler requires neither any external hardware nor any firmware modifications because it leverages Bluetooth's native capability for ad-hoc discoverability and communication, making it suitable for pervasive deployment by simply downloading an app. Each device performs inference online in order to eliminate any privacy issues occurring when transmitting data to third parties. Devices calculate interpersonal distance and relative orientation with respect to each nearby user. Finally, we perform classification on the occurrence of a social interaction, depending on the selected target proximity class.

2.2 Radiolocation enablers

With the necessity of producing geo-referenced data (e.g., in distributed crowd sourcing, user mobility learning, location-based services…), embedded radiolocation capabilities exploiting wireless communications should become common place in most IoT devices and as such, a source of valuable device-centric information (regardless of the device kind, sensor capabilities or enabled services).

2.2.1 Radiolocation technologies

Besides conventional Global Positioning System (GPS) means (yet less and less costly and power greedy), alternative wireless localization solutions have been promoted recently in both WSN and indoor navigation contexts, including low data rate IR-UWB and IEEE 802.15.4/Zigbee radio technologies [10], [11], [12]. Benefiting from low power consumption, the latter also offer appealing peer-to-peer capabilities, which are suitable for mesh and cooperative connectivity in decentralized or versatile networking contexts.

On the one hand, the Impulse Radio – Ultra Wideband (IR-UWB) technology enables Round Trip – Time of Flight and Time (Differences) of Arrival (T(D)oA) estimation with unprecedented timing accuracy, in the order of the nanosecond (i.e. within 30 cm spatial resolution) [13]. Whereas most of the commercialized IR-UWB localization systems already claim metric or sub-metric positioning accuracy, significant progresses have been achieved recently, aiming at better integration, lower power consumption, better compliance with European regulation, and even more aggressive single-link ranging precisions below 10cm (e.g., Decawave [25] and BeSpoon [26]). Side efforts have also been committed to designing synergetic protocols, more notably at the Medium Access Control (MAC) layer level. These approaches enable better support for both ranging and decentralized positioning functionalities. They typically rely on beacon-enabled Time Division Multiple Access (TDMA) superframe structures [14], [15], estimating and compensating harmful relative clock drift effects through the use of cooperative n-way ranging transactions. The RT-ToF gives direct access to the distance between two nodes. On the other hand, various integrated solutions compatible with the IEEE 802.15.4 and Zigbee standards are currently available on the market. All of them can issue RSSI readings at the physical layer for each demodulated packet. Assuming a certain path loss model, such measurements can be exploited for parametric point-to-point range estimation [16]. However, in common environments, the expected precisions of both ranging and positioning protocols are hardly better than several meters [17].

2.2.2 Location information for impersonation prevention and detection

In order to prevent impersonation attacks, conventional cryptographic techniques may not be always suitable in the IoT context, considering the massive deployment of low-cost and low power entities with limited computational capabilities. Furthermore, cryptographic mechanisms usually need a centralized certified entity to distribute, refresh and revoke identity keys and signatures, which may be not necessarily and fully adapted to decentralized networks with temporary ad hoc inter-connections. Alternative non-cryptographic techniques relying on the lower layers of the communication protocol have thus emerged recently [18]. However, software-based methods (e.g., probe request behaviour at the MAC level) or hardware-based solutions (e.g., radiometric fingerprinting, clock skewness or Physically Unclonable Functions –PUF-) are still subject to strong practical limitations, requiring the exchange of numerous challenge-response messages or too specific hardware. They also suffer from residual vulnerabilities (e.g., an attacker can mimic the characteristics of the signatures or rely on machine learning). Additional lower layer techniques are based on the continuous monitoring of physical radio properties such as location-specific Channel State Information (CSI) or RSSI readings to detect identity-based attacks [19], [20], [21]. Channel sounding is then combined with hypothesis testing to determine if prior (authenticated/trusted) and current communications are issued by one unique user. The key challenge is to collect physical radio

metrics over time so as to detect unexpected transients caused by impersonations. In [18], RSSI Similarity based Authentication (SA) and Temporal RSSI Variation Authentication (TRVA) methods are recalled. The SA technique aims at detecting large unexpected RSSI changes between consecutive frames at one receiver. In the following, SA will be used as reference for benchmark purposes. Other contributions aim at protecting, authenticating or validating the location service and data [22], focusing on how to securely establish, disclose or verify the location information itself. High level spatial "anonymization" techniques are also proposed to mitigate the risks in revealing indirectly the user's identity while sending queries to location-based services [23]. Finally, assuming error-free GNSS information in a WSN context [24], each sensor can be uniquely addressed by its 2D coordinates rather than an ID. Given a pre-loaded secret key generated from the initial ID and a system master key, each node securely receives an additional location-based key (LBK) from one mobile entity. Node-to-node authentication and optional pairwise secret key establishment can then be applied using these location-based keys within a pairing-based crypto-system. Security lies in the secrecy of LBK and one can verify that each node has the LBK corresponding to its claimed position for authentication. Therefore, the initial distribution of the LBK to each device from the mobile entity represents a serious vulnerability. Finally, secret key generation from relative localization of a pair of nodes (i.e., the relative distance acting as observed common randomness in a source model of information theoretic secrecy) has been investigated from a theoretical point of view, deriving limits on the achievable secret bit rates, but still regardless of practical considerations on underlying radio technologies and related constraints [29].

2.2.3 Device-centric (secret) information generation

On a side but complementary topic, aiming at the generation of local and unique secrets (e.g., non-deterministic random bit/number generators for headless devices in the context of session key generation or information masking), dedicated hardware (e.g., ring oscillators or thermal noise) or opportunistically diverted embedded sensors such as accelerometers, vibration, magnetic, temperature or pressure sensors can be viewed as sources of entropy. The goal is to make it difficult for an adversary to guess or reproduce the values observed by the legitimate nodes. Recent studies reported in [27] assessing mainly entropy aspects (and focusing more specifically on the pessimistic min-entropy estimator) for a variety of sensors and physical measurements. In particular, it has been shown that the amount of captured entropy strongly depends on the kind of "applicative" embedded physical sensors (e.g., accelerometers are by far more exploitable than temperature sensors). This may be penalizing if too specific/limited devices are used. Most often, the expected theoretical entropy is even overestimated in comparison with that extracted in more practical cases. Thus using multiple sources is recommended, at the price of higher device consumption and complexity. The idea is to provide better information diversity and to continue feeding the entropy pool even when one particular source is corrupted or temporarily abandoned. On the other hand, "default" non-specific wireless metrics such as Link Quality Indicator (LQI), erroneous packets, and maybe less significantly RSSI, which are all obviously location-dependent, have been exhibited as robust entropy sources (against eavesdropping) [28].

2.2.4 Our approach

Overall, techniques reinforcing and/or forging device's identity out of location information have not yet been extended to benefit from the cooperative potential of wireless localization (i.e., the possibility to rely on relative connectivity with respect to neighbours). The heterogeneity of radiolocation modalities (i.e., the possibility to exploit multiple sources and multiple levels of location-based information) has not been really exploited either, although it is expected to provide even better resilience against impersonations, as shown hereafter.

2.3 Localization based on magnetic field

2.3.1 Magnetic field sensing at smartphones

As already pointed out in Sect. 2.2, obtaining precise localization information in indoor environments like buildings is an appealing but still challenging task. Traditional mechanisms such as GPS [45] are not practical inside buildings due to the lack of the signal from satellites. This has resulted in the development of alternative indoor positioning systems with acceptable results, such as those based on WiFi [46] ZigBee [47] and RFID [48]. Nevertheless, a common requirement of these approaches is the deployment of specific devices or additional hardware to be used during the localization process. Consequently, the cost of these solutions is high and frequent maintenance is required. Furthermore, the limited accuracy of indoor localization solutions proposed to date is another issue that needs to be addressed. For example, when wireless technologies are used to solve indoor localization problems, physical obstacles produce signal interferences which affect the performance of the localization systems in question.

As an alternative, the potential of exploiting mobile phones for sensing and context recognition has recently attracted interest from researchers in both industrial [49] and academic communities [50]. This has resulted in a huge range of new solutions for indoor localization [49]. The ubiquitous and longitudinal data that smartphones can provide is expected to revolutionize technological services and spark a new wave of pervasive services.

Contemporary mobile phones contain a number of sensors capable of sensing the user's location and the presence of entities in their proximity. Apart from GPS, which is primarily used for outdoor positioning, GSM, Wi-Fi and Bluetooth signals can also be used for user localization (for extensive readings about ubiquitous localization refer to [51]). On the other hand, integrated sensors in smartphones, such as inertial and ambient sensors, are also being exploited to solve localization problems [52].

2.3.2 Our approach

In this work we present a novel approach for indoor localization based on the magnetometers that are integrated in common smartphones. Unlike most of current phone-based proposals for localization [51], our system does not rely upon an additional support infrastructure. Our solution only requires a personal smartphone able to sense the magnetic field available inside buildings. During a first stage of our system, we generate maps containing the magnetic field profile of the building where the localization problem needs to be solved. This represents the off-line training phase of the system. Then, during the on-line phase, users provide the system with the measurements of the magnetic field vectors sensed by their phone, and using these measurements, our system is able to provide accurate localization data of such users. We evaluate the proposed mechanism based on data samples collected and compare the performance with other existing phone-solutions like WiFi.

Section 3 - High-level scenarios description

The following section presents a detailed analysis of the high-level scenarios in which the different proposed enablers will be employed, in terms of system model, use-cases of interest, involved entities and finally, scenario-specific challenges that need to be solved as part of the research activities to make the enablers suitable to the required context. In particular, the focus is put here on the applicability of such enablers for further demonstration and field validations in the frame of the project.

3.1 Face-to-Face enablers

3.1.1 System model and use cases

One of the fundamental elements of sociability is social interaction. Humans are social beings that interact with each other through verbal or non-verbal communication. Face-to-face (F2F) interaction represents an important dimension of social interactions, which occurs when people are in proximity, maintain mutual-facing directions and exchange verbal signals. They are considered a significant component in the extraction of social signals [60] that people convey in their daily lives. Pentland argues in [61] that location, proximity and signalling behaviour of the user are key properties of human networks that affect the propagation of information (the F2F interaction in this categorisation is understood as signalling behaviour).

In order to study F2F interactions, scientists have mainly used methods such as surveys, questionnaires, human observers, camera recordings. However these techniques lack of automation, which was tackled by the introduction of wearable computing devices. These techniques [31] [33] were able to achieve high accuracy in detecting F2F interactions among people. The incorporation of interpersonal distance estimation with the relative orientation computation of the users to identify the relative spatial arrangement of people, seem to be sufficient for the detection of F2F interactions. Nevertheless, although they were able to manage high inference accuracy, the major disadvantage of these devices is their intrusiveness. People are forced to wear obtrusive devices that are not part of their daily life while they are required to place them on a fixed on-body position e.g. on the chest. An emerging opportunity for overcoming the shortcomings of existing wearable approaches is represented by recent advances in smartphone technologies and opportunistic sensing techniques [30]. The ubiquitous penetration of mobile phones in our society and the widespread use in our daily lives could make them ideal candidates to act as platforms for observation of F2F interactions. The starting point for our research is the question of whether smartphones could serve as suitable platforms for detecting F2F interaction of their human users when carried around by them in their daily life. Current solutions suffer from some limitations: a) inaccurate proximity detection for real world environments, b) disregard the importance of user's relative orientation in a realistically manner, c) lack of efficacious communication mean for exchanging users' facing direction, d) significant delay in sensing and interaction recognition process.

The improved F2F enabler is our enhanced attempt towards realising the vision of detecting F2F interactions through mobile phones by mitigating the above-mentioned shortcomings. It combines opportunistic sensing techniques that exploit readily available sensing and communication capabilities on smartphones into a collaborative sensing system able to recognise F2F interactions. For the F2F interaction recognition, it is able to classify interpersonal distance into different *interaction zones* and combines this with derived knowledge about user's facing directions by computing users' relative orientation and providing similar capabilities as wearable approaches. In order to eliminate the dependence on any external infrastructure and any mobility restrictions while preserving user's privacy, the F2F enabler operates in a completely distributed fashion by detecting

and recording F2F interactions locally on the devices, without requiring a centralised back-end system to perform such inference.

Scenario: Evaluation of the F2F Enabler
The scenario describes the evaluation of the F2F enabler in the context of accuracy, privacy, creation of real-world social graphs and security, as an individual component but also as part of SOCIOTAL framework. It is noted that the scenario and the use cases are defined in D5.1 [8] in the section about the Evaluation of F2F Enabler.

The F2F Enabler utilises only off-the-shelf mobile phones and infers about on-going real-world social interactions. The approach does not require any additional hardware or any firmware modification of the device. It operates in an opportunistic manner without requiring any user intervention in order to perform the inference of any on-going social interactions. A user deploys the SOCIOTAL application, including F2F Enabler and without the need of any particular configuration, the application detects real-world social interactions.

The process of detecting real-world social interactions is based on estimating users' interpersonal distance and their relative orientation. In particular, the interpersonal distance estimation includes the procedure of detecting in which interaction zone the detected user is and also if they are in proximity to interact or not. Estimating the interaction zone of the users may provide valuable information about the social relationship among people [74]. Further, two distinct models based on features of Bluetooth RSSI perform the detection of interpersonal distance. The models were trained by applying machine learning techniques on collected measurements.

The second parameter needed to detect F2F interactions is the relative orientation of the users. To estimate the relative orientation of the users, their current facing direction is required. Each device utilises the walking locomotion of its user and estimates the facing direction [30]. Then, devices exchange the facing directions of their users. This process is performed in the sense of collaborative sensing in an ad-hoc mode. By receiving the facing direction of the nearby device/user the system computes the relative orientation of the two users. Combining the estimation of users' interpersonal distance and relative orientation each device estimates if the users are performing a social interaction.

Knowing about on-going real-world social interactions provides valuable contextual information. In particular, understanding real-world social interactions in long-term monitoring may indicate the social and trust relationship among people. The proposed scenario includes a subject smartphone A that discovers a target smartphone B. Once discovered, smartphone A logs the Bluetooth RSSI of the smartphone B in order to estimate their interpersonal distance. Both devices are computing their users' facing direction. When the appropriate number of Bluetooth RSSI have been collected (i.e., 6 samples) smartphone A initiates a Bluetooth connection between the devices to exchange their facing directions through collaborative sensing. The connection is terminated and both smartphones inform the SOCIOTAL platform about the occurrence of a social interaction. The information about the social interaction may be provided to the Trust Manager for building trust relationships but also to the Authentication component to verify a rule that it has created. Information is shared by means of the SOCIOTAL Context Broker. A context definition containing such information has been provided in deliverable D2.2 [4].

Use case 1: F2F Accuracy Validation
To identify the accuracy of the approach, particular experiments are required. In the evaluation, each user will be provided with a commercial off-the-self smartphone, which has deployed the SOCIOTAL application. Each user will be asked to stand for a few seconds and then will start to walk to estimate

the facing direction. The user will continue the daily routine and during that period the detected social interactions will be logged. In order to recognise the ground truth, once a social interaction is detected, a mobile notification will be prompt to the user; the frequency of notification will be configurable through the application. Thus, the user will verify which social interactions were correctly detected. The contextual information logged by the application will be forwarded to the SOCIOTAL platform for additional processing, as previously described.

Use case 2: F2F Privacy Level Validation

Understanding the privacy level of the approach is important. The F2F Enabler performs the inference of on-going social interactions on the device and does not transmit sensed data to third party components. Only the result of the inference is transferred to Trust Manager and Authentication component, to understand the level of trust of the users and to verify created rules, respectively. To validate the privacy of the approach, a tool proposed in the User Empowerment for Enhanced online Management Project (USEMP) [64] will be utilised. Using the APIs provided by this USEMP platform, the developed tool provides any SocIoTal F2F application user with a daily report about various privacy aspects (typically, regarding the collected and shared information) and collects also some feedback about the perceived privacy of the application. In detail, a subject smartphone A creates this daily report and receives directly feedback from the user. To further process the daily reports and users' feedback the information will be transferred to SOCIOTAL server.

Use case 3: Creation of Real-world Social Graphs

Detecting F2F interactions allows the creation of real-world social graphs. Social networks prompt the users to classify their relationships with other people. F2F enabler allows, through identifying social interactions and social relationships, the creation of real-world social graphs. In parallel, users will be prompt with their detected real-world social graph, giving them the opportunity to validate the detected graphs. Further, information acquired from users' social networks (such as Facebook, Twitter, LinkedIn etc.) will be compared with the real-world social graph that SOCIOTAL application created. In particular, a subject smartphone A detects social interaction with target smartphone B. This leads to the creation of an edge on the real-world social graph. This information is forwarded to the SOCIOTAL server.

Use case 4: F2F Security Level Validation

The F2F enabler operates in a distributed manner and communicates also with SOCIOTAL platform to provide it with results of the inferences. In essence, it constitutes a mobile and connected device that could be vulnerable to malicious users. To understand the effect of malicious users, there is a need to validate the security of the component and the system. Measuring the unauthorised social relations may provide an indication of the security combined with the security analysis and the KPIs of the D5.1 [8]. In particular, a subject smartphone A detects a social interaction with target smartphone B. However, the detected social interaction is erroneous due to either maliciously injected sensor data or because of the compromised communication channel between the device A and SOCIOTAL platform.

3.1.2 Involved actors and technology

In the particular scenario described above, the involved actors are focused on the smartphones carried by people-users that participate to the enabler evaluation. Assuming the aforementioned scenario, a subject device (smartphone A) discovers and estimates if its user is performing a F2F interaction with the user of a target device (smartphone B). In this particular deliverable we will focus on the evaluation of the enabler in the scope described as "Use Case 1", while a more comprehensive evaluation including the other mentioned use cases, will be provided as part of WP5 work and reported in D5.2 [9]. This is in line with the objective of WP3 and this deliverable, to

provide a stable implementation of this and other enablers to be used in following trials and pilot experimentation.

The F2F enabler utilises Bluetooth interface of off-the-shelf smartphones to discovery nearby devices. Once a device has been discovered, the enabler logs the Bluetooth RSSI to estimate interpersonal distance between the two devices and further between their users. Interpersonal distance refers to the ability to estimate if the users are in proximity or not, and in which interaction zone they are. In parallel, leveraging user's walking locomotion, each device computes the facing direction of the user. A collaborative sensing technique, allows the exchange between the users of contextual information such as interpersonal distance and facing direction. Once each device has acquired the facing direction of the opposite device, it is able to calculate the relative orientation of the torsos of the two users. Having estimated the interpersonal distance and the relative orientation of the users, the devices are able to recognise if the users are performing a F2F interaction or not.

Further, detected F2F interaction by the enabler will provide valuable information to the Trust Manager and the Authentication component. In detail, F2F interaction followed by contextual information may indicate the social relationship and levels of trust among people. Also, the F2F enabler may validate a rule created by the Authentication component.

3.1.3 Scenario-specific challenges

The main challenge of the F2F enabler is to be able to detect F2F interaction in an opportunistic, unobtrusive and accurate manner. First, the enabler should not require the user participation in order to perform the inference about existing social interactions. Secondly, the F2F inference should be performed accurately allowing the system to trust the outcome of each device.

Another challenge is the fact of detecting F2F interactions in a privacy preserving way, as smartphones with deployed the F2F enabler; log contextual information, which should not be exposed to others. Extracting the appropriate contextual information from detected F2F interaction to create social graphs constitutes also a significant challenge. Finally, ensuring the operation of the F2F enabler with the SOCIOTAL core system in a secure way is important for the correct and reliable outcome of the component. Mechanisms for securing the system are required to prevent malicious users to falsify information such as identities, inference and relationship data etc. The incorporation of SOCIOTAL Trust Manager may comprise a valuable defending mechanism.

3.2 Identity reinforcement and impersonation prevention in Bubbles of Trust

The radiolocation enablers provide context information about end-devices that can be used (directly or indirectly) in established Bubbles of Trust for identity management, trust evaluation or more simply, for general security purposes (e.g., assisting authorization, authentication or secure communication procedures). In this section, we focus mainly on high-level scenarios that would make sense for possible integration into the SocIoTal framework (even if not necessarily implemented/involved in final physical demonstrations), in compliance with the specified communication and management entities. Further details are provided in Sect. 8.2.

3.2.1 System model and use cases

In the most generic case, input information can comprise absolute 2D coordinates and/or raw intermediary measurements (i.e., single-link metrics used as inputs for the estimation of 2D coordinates). For instance, positions can be issued by embedded GPS sensors in outdoor. Alternatively in indoor environments, they can be also computed through WiFi RSSI-based fingerprinting with respect to surrounding Access Points (APs) or using Zigbee RSSI, Bluetooth-LE RSSI or IR-UWB RT-ToF measurements with respect to fixed Wireless Sensor Network (WSN) anchors and/or mobile neighbouring devices (respectively through trilateration and cooperative positioning

algorithms). Ultimately, device-specific and link-specific location-based pseudonyms (LBP) are produced out of the previous radiolocation sources. This means neither that the devices will have to host all these heterogeneous means (i.e., GPS, WiFi, LDR radios) nor that the latter will be operating simultaneously (e.g., one may want to save power consumption or limit interferences). But depending on the security strategy (i.e., given some risks assessment policy in a particular environment) and on the intended usage of the output information (e.g., for identity management, trust and reputation management, simple "protection" overlay assisting authentication procedures…), only the most adequate subset (among available radiolocation inputs) might be combined for pseudonym generation.

As already pointed out in D3.1.1 [5], possible applications concern static Wireless Sensor Networks (WSN) (e.g., for house automation, equipped elements of outdoor furniture in smart cities…). When the operator introduces a novel node/device, the latter is indeed uniquely specified by its physical insertion into the environment. But such enablers are mostly intended in local and spontaneous groups of mobile users, formed in an "ad hoc" way. These users are physically co-located (e.g., in an office room…) and presumably static during the LBP establishment process. Practically speaking, this shall just imply that the users restrict their movements for the duration when the location-based pseudonyms are generated, given a certain quantization resolution of the input location parameters (i.e., if the quantization procedure is relaxed enough relatively to the observed dynamics, so that slight modifications of the perceived/measured physical location may lead to the same pseudonyms as that generated in a purely static case). Hereafter follow more concrete scenarios and use cases.

Scenario 1: Trust Assessment through Location-Based Pseudonym Verification

First of all, raw radiolocation measurements (i.e., seen as contextual information issued from embedded sensors) and their derived location-based pseudonyms (i.e., higher-level contextual information elaborated out of these raw radiolocation measurements) can be exploited for the prevention and detection of impersonation attacks, and more generally, for trust assessment. Classically, once authentication/authorization procedures (assisted by LBP verifications) are completed for one given device, the latter could have access to restricted resources, or reciprocally, one verifying tier could have access to the resources of the authenticated/authorized device.

Use Case 1a: Prevention of Impersonation Attacks with Link-Dependent Pseudonyms

In a first concrete example, devices (e.g., smartphones) are assumed to be endowed with Low Data Rate (LDR) communication capabilities, enabling direct peer-to-peer (P2P) links and ranging (e.g., IR-UWB, Bluetooth-LE or Zigbee radio modems, seen as embedded sensors, like in commercially available *SpoonPhones* [26]). Given two particular devices, namely A and B, which aim at building (or maintaining) a trust relationship, A could provide "continuous guarantees" to B about its identity and trustworthiness in a steady-state regime (i.e., once B has preliminarily authenticated A based on conventional procedures). For this sake, A is expected to generate a unique location-based pseudonym (LBP) out of some location-based radio measurements that it has performed with respect to B (e.g., P2P RT-ToFs, possibly along with relative clock drifts or P2P RSSIs). The goal of Device B is then to i) get directly the range-based pseudonym claimed by A (preferably in a secure way) without making any query to a centralized entity, and finally to ii) compare this pseudonym with a local corresponding guess (i.e., generated by B out of its own radiolocation measurements with respect to A). This is a dynamic process in the sense both devices A and B can physically move and thus, will practically change the value of both A's LBP and related B's guess as a function of time. Moreover, after disclosing its LBP, one device should re-compute it to prevent eavesdropping attacks. Optionally, beyond simple verification aspects, the latest available LBP, respectively generated and guessed by the two legitimate piers (i.e., at the previous time stamp) but not yet disclosed, could be used to "secure" the current interrogation phase from B to A while getting the

new LBP claimed by A (i.e., the public ID of A being substituted by its latest available LBP during the direct transactions with respect to B and/or used as session key to cypher the transaction using symmetric cryptography). Note that time stamps associated with the intermediary measurements and/or with the final location-based pseudonym shall be optionally collected for further validity check (i.e., compared with a pre-defined validity timeout). This continuous LBP verification procedure can then assist and complete other context-based trust assessment procedures relying e.g., on location/proximity information verification or token-based authorization verification (as shown within other enablers). In Sect. 8.2, we will provide further details about the data flow and the concerned SocioTal management blocks in this example.

Use Case 1b: Prevention of Impersonation Attacks with Position-Dependent Pseudonyms

In this other example, the goal of B is still to verify the LBP claimed by A, but now with the corresponding LBP stored as contextual information (i.e., about A) by a centralized entity. Moreover, A and B do not have necessarily to perform mutual radiolocation measurements with respect to each other. So as to feed the on-device pseudonym generation algorithm, we assume that A is just enabled with absolute localization means (GPS, WiFi fingerprinting with respect to APs, trilateration with respect to fixed WSN anchors). However this positional information could still be advantageously complemented by relative information with respect to neighbours (but again, not necessary with respect to B). Similarly to the previous use case, time stamps associated with intermediary measurements and final pseudonyms may be used for further validity check. Practically, A will first sense its own radiolocation measurements and deliver a local position estimate (or alternatively get the result from an external localization service). All this input information is then used to elaborate Device A's LBP. B will require directly the (claimed) LBP from A and compare it with the LBP retrieved from the centralized entity (e.g., IoT context broker), as previously published by the authenticated A. In this case again, in a steady-state regime, this LBP verification procedure (carried out at B about A) can assist e.g., further trust assessment or token-based authorization procedures.

One practical synopsis showing the use of such LBP incorporating both the location and the relative connectivity in dynamic authentication procedures during discovery is illustrated below:

1. As Device B knows its own location (e.g., from radiolocation means or alternatively, magnetic field sensing according to another device-centric enabler) and since it starts "(re-)exploring" a new place, it asks a centralized entity (e.g., database on the back-end side e.g., IoT context broker) to provide the list of neighbouring devices in that place (i.e., devices sharing the same geographical area, presumably in range of direct communications and already authenticated with respect to that centralized entity);
2. Based on Device B's location, the centralized entity selects the public IDs of the most plausible devices as potential neighbours for B, along with their LBPs and time-stamps
3. The centralized entity sends the required information back to B in a secure way.
4. After retrieving the list of expected neighbours, Device B then broadcasts a query containing at least its own public ID to the present devices around, asking them directly to disclose their "claimed" LBPs.
5. If in range, Device A detects Device B's query packet, and sends back a response, which contains at least its LBP. The latter should reflect its relative connectivity with respect to its own neighbours (i.e., the connectivity experienced till B arrives), along with its public ID;
6. Device B then receives the response from A and verifies that the claimed LBP is compatible with that retrieved (still about A) from the centralized entity, following a one-shot verification approach (i.e., the LBP is not more valid after being disclosed once).
7. Right after sending its response to B, A must renew its LBP based on the refreshed connectivity information (now including B, by incorporating e.g., one more relative RSSI-based ranging measurement with respect to B in additional to its measurements with respect to other "established" neighbours);

8. Device B then pushes its new LBP, along with its absolute 2D location and the associated time-stamp into the centralized database on the back-end side (e.g., IoT context broker) in a secure way;
9. Device A monitoring continuously its location, it can ask again the centralized entity to provide the updated list of LBPs associated with physically present neighbours, and jump back to Step 2 above for continuous LBP check (regardless of devices physical mobility).

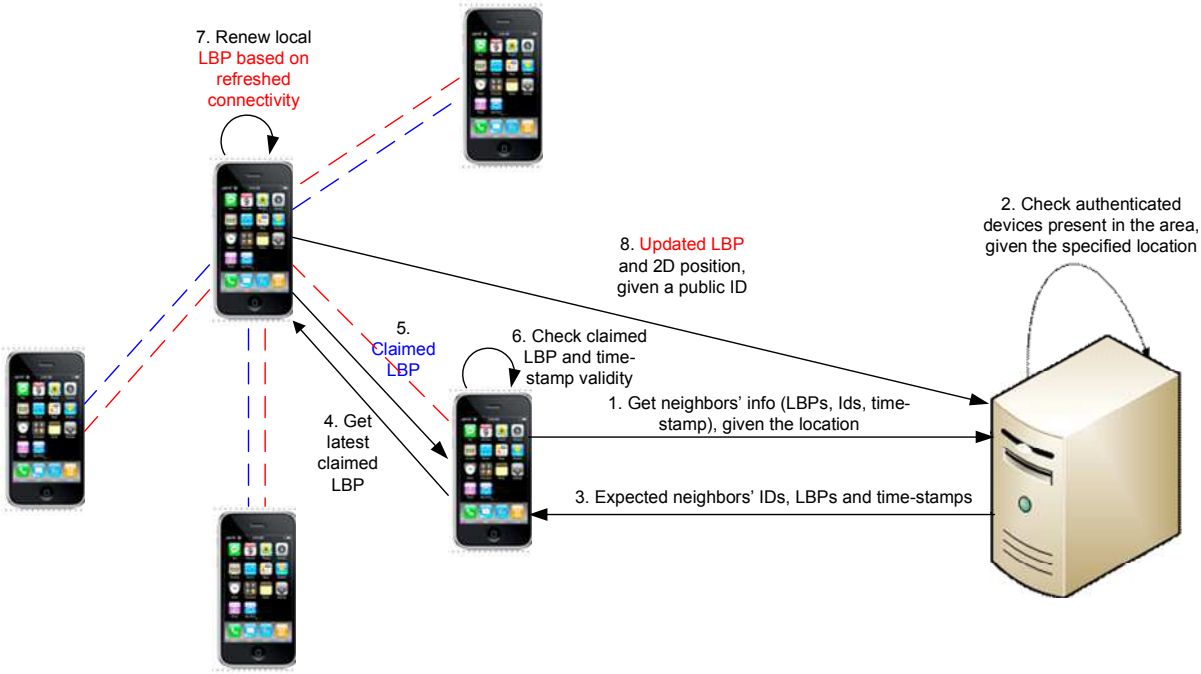

Figure 3. Ex. of dynamic position–based LBP verification for reinforced authentication

Note that it is supposed that both A and B have been preliminarily authenticated with respect to the centralized entity (each independently) and that they can always publish / retrieve information with respect to this centralized entity in a secure/private way.
The LBP being disclosed only once (immediately renewed to prevent eavesdropping) and integrating hardly predictable information (including relative connectivity conditions w.r.t an arbitrary large and a priori unknown number of neighbours), impersonation attacks based on direct LBP guess or message interceptions/manipulations (from/to the database and between devices directly simultaneously) are thus made significantly complex and costly.

Scenario 2: Secure/Private Communications with Location-Based Pseudonym

Once a trust relationship has been established (possibly triggered by face-to-face enablers) between two devices, they could mutually reinforce their identity profiles to share even more securely their personal data in the near future, directly based on pseudonyms.

Use Case 2a: Secure/Private Communications with Link-Dependent Pseudonyms

Device A can use its own LBP (i.e., the LBP it has locally generated with respect to B) to temporarily substitute its public ID in further data exchanges so that device B (e.g., while renewing the LBP at each new transaction to prevent from eavesdropping attacks), based on its own guess about Device A's LBP, can thus use the same substituted ID for A in both Tx (i.e., to address the right neighbour while transmitting information) and Rx modes (i.e., detecting and selecting the right transmitting

neighbour while receiving information). However this concept is pretty close from the LBP verification solution presented previously for authentication purposes.
But Device B can also use its own guess about A's LBP to protect data exchanges with A (e.g., using A's LBP as a common private key in symmetric cryptography or as session key). In this case, the LBP approach does not present any specific advantage in comparison with other conventional random secret generators but it mostly offers another alternative in case such generator is not embedded/available/working properly at the device.

Use Case 2b: Secure/Private Communications with Position-Dependent Pseudonyms
This use case is the same as previously (Use Case 2a), including ID temporary substitution or secret generation (as cryptographic seed), all except but A's LBP is now position-dependent and the information (locally generated at A) is also stored in the back- end so that any authenticated device B with respect to the back-end can require and get A's LBP to protect further transactions and communications with respect to A.

Scenario 3: Explicit Radiolocation Information for Mobility Learning

Here, the acquired radiolocation information (e.g., time-stamped 2D positions) can be stored and further processed (either locally at the device -at least in part- and/or in a centralized way) to learn, classify and predict usual mobility habits/patterns of users or devices (in the arbitrary short/medium/long terms), as well as to detect and assess unexpected mobility behaviours, in both conservative (i.e., security-oriented) or positively incentive ways. Since this scenario and all the possible related use cases (e.g., in data crowd sourcing application for smart environments) are not directly treated at the device-specific enabler level but mostly handled by the Reputation and Trust manager, it will be not detailed hereafter.

3.2.2 Involved actors and technology

The actors and technologies involved for the radiolocation enabler are as follows:

- Users equipped with multi-standard mobile phones, enabled with:
  - Absolute positioning means (e.g., GPS, positioning w.r.t. localization infrastructure such as IR-UWB or Zigbee/IEEE 802.15.4 w.r.t. anchors or WiFi APs or RTLS BSs) and/or…
  - Relative peer-to-peer ranging or cooperative positioning means (RT-TOF over IR-UWB links, RSSI over Zigbee/IEEE 802.15.4 links or WiFi direct links)
- WSN nodes, enabled with:
  - Absolute positioning means (e.g., GPS, positioning w.r.t. localization infrastructure such as IR-UWB or Zigbee/IEEE 802.15.4 w.r.t. anchors or WiFi APs or RTLS BSs) and/or…
  - Relative peer-to-peer ranging or cooperative positioning means (RT-TOF over IR-UWB links, RSSI over Zigbee/IEEE 802.15.4 links or WiFi direct links)

It is worth noting that the radio medium used for the wireless localization functionality does not necessarily coincide with that used for the secured communications or direct data/context retrieval from a neighbouring device (e.g., considering smartphones endowed with IR-UWB for precise localization or peer-to-peer ranging and BT-LE or WiFi direct for peer-to-peer data communications).

3.2.3 Scenario-specific challenges

The main challenges related to the intended use cases are as follows:

- Degraded precision of the input localization service in typical indoor environments (e.g., Non Line of Sight (NLoS) with respect to fixed APs/BSs/anchors or fellow "mobiles", imprecision/availability of model parameters, for instance for RSSI-based ranging);

- Underlying/supporting localization functionalities possibly subject to their own security/privacy issues, such as passive eavesdropping, service denial or disruption (e.g., due to the presence of location-specific traffic on public over-the-air channels);
- Mobility support for the perennial tracking/management of new generated location-based IDs;
- Generating unique IDs conditioned on graph/deployment configuration is not always guaranteed (in terms of graph rigidity, nodes identifiability and localizability);
- Attacks based on extensive brute-force guesses (regarding deterministic location-based IDs) are still hardly avoidable.
- Most schemes are not intended from scratch but necessitate the existence of an initial authentication procedure between devices.

3.3 Indoor positioning enablers based on smartphones

3.3.1 System model and use cases

Given the indispensable role of mobile phones in everyday life, phone-centric sensing systems are ideal candidates for ubiquitous observation purposes. The increasing development of wireless communication technologies and ambient intelligence is enabling a seamless integration of smart objects in our everyday lives. This emerging trend, along with the global deployment of mobile devices, such as smartphones or tablets, are redefining the way people exchange information and communicate among them as well as with their surrounding environment, transforming current physical spaces into smart buildings. These incipient ecosystems are expected to be composed by sensors, smart devices and appliances that can be remotely monitored and accessed by users or cloud services, resulting in a new generation of intelligent and ubiquitous environments.

However, unlike the current Internet, the realization of these scenarios requires higher security and access control restrictions, as well as sophisticated mechanisms to calculate levels of trust and reputation of users.

On the one hand, physical objects in typical buildings are being integrated into the Internet infrastructure with network and processing abilities, making them vulnerable to attacks and abuse. On the other hand, in these pervasive scenarios, services and resources can be accessed via mobile devices anytime and anywhere by common users, and users can decide to share their data and objects with other users according to their levels of trust and users reputation.

While this trend provides significant benefits regarding availability of services and sharing of information, there are everyday situations in which users could abuse these services if location data is not considered at security level.

For example, considering location-aware mechanisms for trust evaluation, in the smart buildings context, a user located at a certain room may like to share his data only with users located in his same location. In this way, a specific level of trust can be automatically established among people located in the same room, because all of them can be seen as belonging to the same ecosystem. For example, we can think in a classroom with students and one of them wants to share his class notes with the rest of students of such room.

The correctness and effectiveness of location-aware security mechanisms is closely related to the accuracy of the location information and the definition of security zones, that is, the area where security aspects like access control, trust, reputation, etc. may be established or rechecked. However, in the context of smart buildings, how this location information is obtained is a challenging task since traditional mechanisms such as GPS are not useful due to the lack of signal in indoor environments. This has resulted in the development of alternative positioning systems, as those based on WiFi, ZigBee and RFID, with acceptable results in buildings. Nevertheless, a common

feature of these approaches is the need to deploy additional hardware, consequently, the cost of these solutions is high and frequent maintenance is required. Additionally, an inherent aspect of localization systems is the limited accuracy of the location information, due to physical obstacles or interferences. Therefore, in location-aware trust mechanisms, authorization decisions should properly consider this degree of uncertainty associated with the vagueness of localization systems.

In order to overcome the aforementioned challenges, the localization mechanism for smart buildings is able to provide accurate data to be included in security aspects of smart services. The proposed indoor localization system is based on the use of sensors which are integrated in common smartphones.

Different use cases in the scenario of indoor environment are described below, all of them integrating localization data estimated by our indoor localization mechanism based on magnetic field.

Scenario: Location-Aware Access Control for Indoor Environments

The proposed scenario shows how the SocIoTal Capability based Access Control system can rely on the Indoor location enabler to make authorization decisions accordingly. The scenario is directly related to the trial defined in deliverable D5.1 [8] about Evaluating the Location-aware Access Control for indoor environments.
The indoor localization solution uses the smartphone built-in magnetic sensors to make security mechanisms totally independent on the type of devices and available signals in buildings. The sensed magnetic field is a combination of the effects of the Earth's magnetic field and that of surrounding objects. The effect of surrounding objects can be divided into deterministic interference, which includes the effect of ferrous materials (soft iron) and magnetized materials (hard iron), and non-deterministic interference. The effect of nearby objects can distort or even drown out the weaker direction of the Earth's magnetic field for navigation and localization purposes in buildings.

A methodological approach is used to generate the buildings maps containing the magnetic field distribution used as map of fingerprints. Then, based on such maps, location estimations are calculated using a combination of Radial Basis Functions Networks and Particles Filters.

Using the magnetic field maps of the buildings, an indoor location service is responsible for computing the position of a device inside the building. In this way, devices can ask this service in order to get the distance where a requester user is placed when trying to access to their services; consequently, certain services can be only provided when users are placed inside the authorization zone of some smart objects.

The position worked out by the Indoor location service can be taken into account to make authorization decisions accordingly. The Figure 4 below depicts the proposed scenario to perform location-aware access control in indoor environments. The smartphone A acting as a subject requests to get access a resource being provided by the smart device B. Before allowing it to access to his resource, the target device B evaluates both the capability token as well as the A's position, which must be located inside B's security zone. The context that determines the smart object B position comes from the indoor localization enabler, which relies on magnetic field measurements.

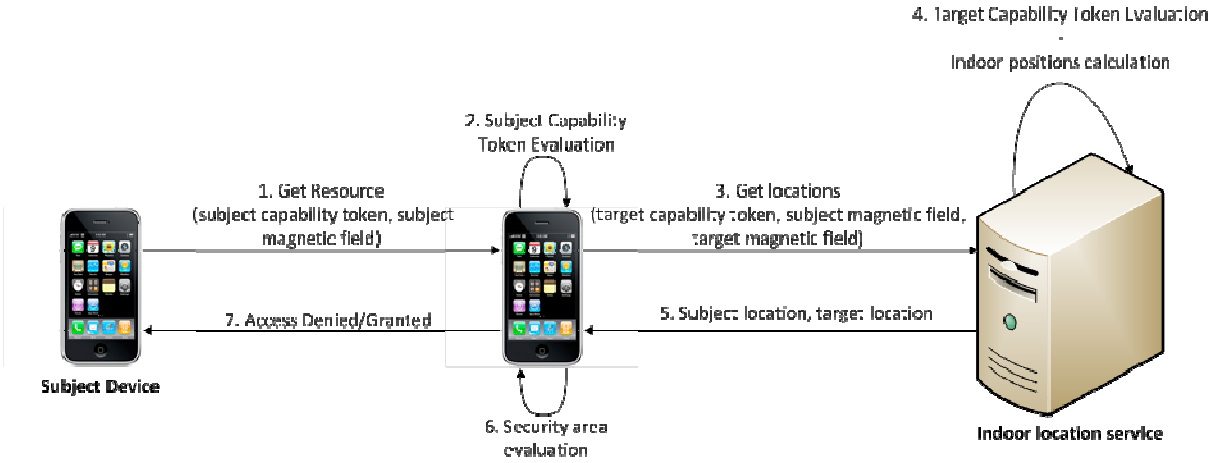

Figure 4. Location-aware access control for indoor environments

Use case 1: Location based Access Control

Firstly, the use case requires an offline stage where the smartphone of user A contacts with the Authorization Manager in order to get an authorization credential (i.e., a capability token) to get access to smart objects. Notice that this phase requires the authentication process. Once the smart object A is successfully authenticated, the Authorization Manager evaluates the policies and generates (if allowed) a capability token with the set of privileges associated to the smart object. Figure 4 depicts the main interactions of the use case. Once the smartphone has the token, a smartphone A acting as subject device wants to make use of a resource hosted by device B. The smartphone A uses its capability token to present it against device B, which validates the token (signature, grants, etc.) and checks A's position against a localization service, since only those devices located near to B are allowed to get access.

Use case 2: Indoor Location based on Magnetic Measurements

The Indoor location service analyses the magnetic field measurements coming from the device to come up with the location of the subject device. Then the location position is send back to the device B that checks the token and the position and makes an authorization decision. In order to obtain the location the Indoor location service uses the generated buildings maps containing the magnetic field distribution.

3.3.2 Involved actors and technology

The subject device (Smartphone A) is a common smartphone that perform a request against the smartphone B. The target device (Smartphone B) relies on the Indoor location service to obtain the subject location.

The Indoor location Service analyses the magnetic field measurements coming from the device to come up with the location of the subject device. The service also evaluates the capability token.

Optionally, a Trust Manager component can be used in this scenario to quantify trustworthiness of the subject device. The Trust manager can be deployed either in the same location of the Indoor location service or separately.

3.3.3 Scenario-specific challenges

The usage of the magnetic field mechanism to obtain the indoor position of the device arises some challenges. The first challenge is related to the localization quantification, and in turn, the accuracy of the magnetic measurement. In this regard, the system requires a previous learning stage very important to calibrate the localization enabler.

Security concerns are also important in this scenario, since the subject could fake the magnetic measurement during the enabler operation. Thus, another challenge is about quantifying the subject's trustworthiness to be sure that the entity is trusted enough, and therefore it is not likely for him to act maliciously against the Indoor location service or the target device. To avoid this situation, both the target device and the indoor localization service could request the SocIoTal trust manager component in order to guarantee that the subject is trusted enough.

Section 4 - Face-to-face enablers

4.1 Improvements to the nominal embodiment

One of the fundamental elements of sociability is social interaction. Humans are social beings that interact with each other through verbal or non-verbal communication. Real- world social interactions occur when people are in proximity, maintain mutual-facing directions and exchange verbal signals. Further, they are considered to be a significant component of social signals [60] that people exchange in daily life. Pentland argues in [61] that location, proximity and signaling behavior of the user are key properties of human networks that affect the propagation of information.[1] Thus, providing an opportunistic and unobtrusive method to quantify users' social interactions would constitute an important step in human behavior understanding for several fields including social sciences.

In D3.1.1 [5], we presented an accurate and reliable system, named DARSIS, for measuring social interactions opportunistically based on off-the-shelf smartphones, without the need of any external hardware. First, we developed a novel machine-learning based methodology for estimating interpersonal distance among users with only 6 Bluetooth RSSI samples. We showcased a model for inferring if users are in proximity or not. Second, we incorporated into the social interaction detection process, a method for computing the relative orientation of a user that allows estimations to be performed regardless of the on-body wearing position. Third, we introduced a collaborative sensing mechanism allowing devices to exchange sensed information such as user's facing direction and Bluetooth RSSI measurements. These components were incorporated into a coherent system, enabling accurate and pervasive sensing of real-world social interactions.

Proxemics have preoccupied the field of psychology due to importance of interpersonal distance among people in order to characterize their relation. In [74], Hall mapped interpersonal space among people into four interaction zones based on their relationship: a) Public, b) Social c) Personal d) Intimate. To detect these interaction zones, an accurate interpersonal distance estimation technique is required. As argued in Section 2, we focus on RSSI-based techniques to estimate distance. However, current approaches lack of accuracy because of the fluctuation observed in RSSI and are dependent on numerous samples e.g., 20 samples [44]. To tackle these problems, in next two subsections we present both architecture and implementation details regarding the training process for developing models that accurately classify interaction zones and proximity.

4.2 Architecture and implementation updates

The main novelty with respect to the previous enabler internal architecture is the incorporation of an additional module for interaction zone detection. This component is placed in the pre-processing in Figure 5. Previously the enabler was able to infer only if the user was in proximity or not through the proximity component. Thus, we added a new component for interaction zone detection.

[1] Social interaction in this context is interpreted as signalling behaviour

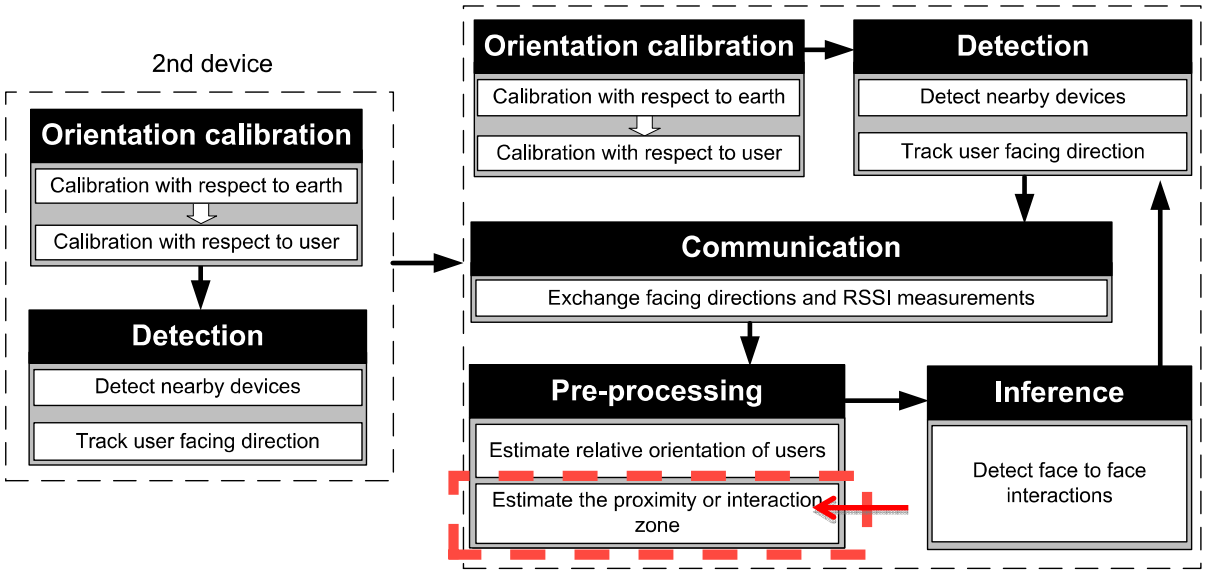

Figure 5. Architecture of F2F enabler

Given the lack of accuracy in distance estimation through empirical evaluation in RSSI-based state-of-the-art techniques, we developed a new machine learning based solution given only a 6-sample RSSI window. The key idea is to perform classification of interaction zones in a hierarchical fashion to achieve high accuracy while requiring only a few RSSI-samples allowing the detection of short- time interactions. Our DARSIS Hierarchical Classifier (DHC) consists of 2 layers of machine learning models. Classifiers at first layer are particularly trained for a given region i.e., public, social or personal zone; a set of domain expert classifiers. Classification confidence values from this layer provide a fuzzy membership for each window of RSSI values. The second layer classifier strives to find optimum thresholds for mapping membership values into the correct interaction zone. It is worth noting that all the classifiers we evaluated faced significant difficulties in differentiating between personal and intimate interaction zone with adequate performance. We therefore merged intimate zone, and focused our classification on personal, social and public zones.

In addition, we develop DARSIS Proximity Classifier (DPC) to detect if user is in proximity or not. To mitigate problems associated with previous similar works, particular consideration is made in training the model and selecting the input feature sets.

Our approach for developing interpersonal distance classifiers is to first generate a generic training set and then produce a bank of features from this training set. Next, we perform feature selection to select the most informative and less redundant feature set. Finally, we evaluate different classifiers at each level of hierarchy to find the most appropriate selection.

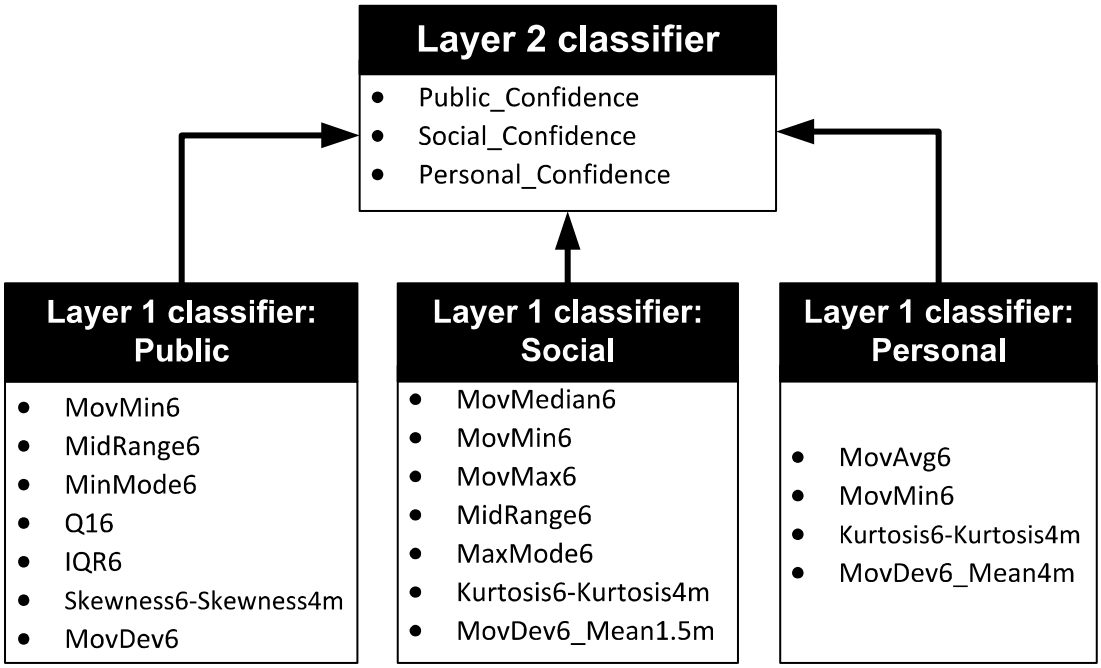

Figure 6. Features for 2-Layer DHC

In order to construct the training set for our classifiers, we conducted a data collection campaign in an indoor office environment using HTC One S smartphones. During each experiment, the Bluetooth interface of one device was set to discoverable mode, while the other device was performing a discovery. The experiment acquired RSSI readings between two smartphones for eight distinct distances from 0.5m to 4.0m, every 0.5m. For each different distance, three device orientations were scrutinized: a) Screen-to-screen b) Screen-to-Back c) Back-to-Back. After several experiments we concluded that these three above vertical orientations are representative for the effect of facing direction change. For reasons of statistical significance, a large number of measurements (2000 samples) was acquired for each distance/orientation combination, producing in overall 48000 RSSI measurements. Due to extremely lengthy data collection process, we replaced humans with water-filled cylinder [68] to which devices where attached. The devices were always in vertical orientation and were placed at a height of 0.8m from the floor, to replicate a common wearing position i.e., trousers pocket [69].

Based on this dataset and given a maximum window of 6 Bluetooth RSSI samples, we created a large feature set (3050 features) including several statistics. Initially a subset of features is selected by considering the level of consistency [70] of each distinct feature with respect to the target class. After obtaining this subset of features, a wrapper subset evaluation [71] was performed to acquire an optimised feature set for the selected classifiers.

We initially trained DARSIS Single Classifier (DSC) for all three zones based on several algorithms. Through evaluations we determined that MultiBoostAB [72] with decision tree J48 [73] was the best classifier for our models. It achieved the highest accuracy in a robust manner due to its native capability for variance and bias reduction. In order to boost its performance further, we explored the use of a DHC. The previously achieved accuracy and inference robustness of MultiBoostAB, led us also to train the models of each layer of the DHC based on the same algorithm, which succeeded in an even higher accuracy than the DSC.

When people are in social or personal zones they are in proximity to interact. In this context, in addition to our interaction zones classifiers, a DPC is also developed that infers if the users are in

proximity or not. In particular, our DHC could detect proximity or not through discriminating between public and other zones. Thus, personal and social interaction zones of DHC are classified as proximity.

Section 5 - Radiolocation enablers

5.1 Improvements to the nominal embodiment

In comparison with the description already provided in D3.1.1 [5], one first improvement regarding the radiolocation enabler lies in the possibly to use explicit P2P RSSI reading as input measurements so as to feed concrete range estimators before quantization (e.g., according to [16]), whereas practical implementations involving IR-UWB RT-ToF input measurements have been mostly considered and discussed so far. This tends to make the proposal less specific and more flexible (in terms of addressed scenarios and implementation potential), considering that most of mobile devices endowed with direct wireless communication capabilities can already produce RSSI readings by default today. This is also partly compliant (and to some extent converging) with the local connectivity table required in one particular embodiment of the F2F enabler described before. Moreover, besides the LBP generation and distant (legitimate) guess considerations already treated in D3.1.1 [5], concrete cross-verification LBP schemes relying on the SocIoTal context model have been put forward, as defined in Sect. 3.2 and 8.2. Besides, no particular performance improvement is expected in comparison with the material provided in D3.1.1 [5]. The reader is thus invited to refer to the latter document for further details about architecture and implementation aspects (i.e., beyond the simple reminder below).

5.2 Architecture and implementation updates

In comparison with D3.1.1 [5], the overall architecture and implementation of the radiolocation enabler is unchanged. Figure 7 recalls the information flow for the realization of both location-based pseudo generation and detection of impersonation attacks (including preliminary wireless connectivity discovery and location information acquisition before LBP generation and verification).

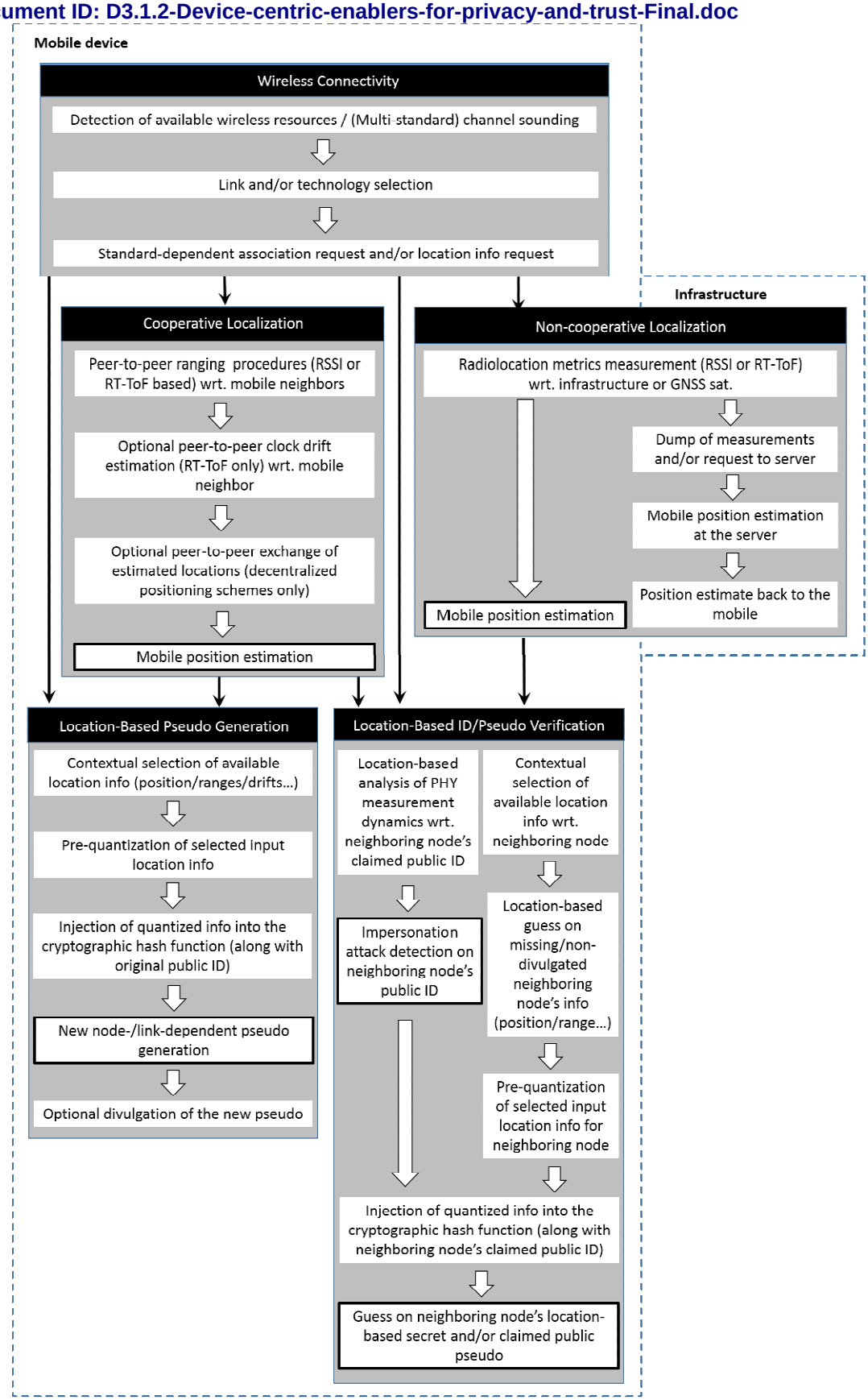

Figure 7. Information flow between building sub-components of the radiolocation enabler.

In the most generic case, if multiple sources of location dependent information are available at a given device, one can perform heterogeneous data integration, as follows:

$$PS_i = f(ID_i, [x_i, y_i], \{d_{ij}\}_{j \in Ne(i)}, \{\gamma_{ij}\}_{j \in Ne(i)} \ldots) \quad (1)$$

where $[x_i, y_i]$ are the 2D absolute Euclidean coordinates of node i delivered by, e.g., GNSS or a WiFi-based localization system, $\{d_{ij}\}_{j \in Ne(i)}$ is the set of measured peer-to-peer distances w.r.t. neighbours $j \in Ne(i)$ (possibly drift-biased, with relative clock drifts $\{\gamma_{ij}\}_{j \in Ne(i)}$), and $f(.)$ is an arbitrary (public) hash function (e.g., SHA-1) to avoid collisions of pseudonyms between legitimate nodes.

Note that simpler input metrics (still accounting for the "physical insertion" of the device) could be considered as well, such as node i's (N-hop) connectivity table w.r.t. to nodes $j \in Ne(i)$.

The generated pseudonyms between different legitimate nodes can be shared/retrieved through direct communications during a "secure" short time period or by inference (i.e., the neighbours of a given node try to guess its pseudonym to the best of their knowledge). In the inference case, the measurements feeding the hash function are quantized in order to provide better stability and reciprocity against input measurement noise for legitimate nodes. The setting of the quantization step is a key parameter that can be adapted online according to empirical measurement statistics. It is obvious that inference is preferable for security reasons.

However, other alternatives exist, for instance by securely publishing and storing the locally generated LBP as context information in to a centralized entity (e.g., in a context broker on the back-end side) and then retrieving this information after secure query to this very centralized entity.

Regarding LBP generation, two preferred embodiments are possible:

- ***One global pseudonym per node***: the hash input can contain location information (e.g., position, relative distances w.r.t. all the neighbors) and/or connectivity information (e.g., adjacency information), and/or further device-dependent information (e.g., relative clock drifts w.r.t. all the neighbors). Inferring this type of pseudonym requires information exchanges, which, in the long term, could lead to a disclosure of the network characteristics and can limit the advantage with respect to attackers. In this case, the pseudonyms must be shared securely.
- ***Link-dependent pseudonyms***: the hash function outputs a new ID of the present node w.r.t. each neighbor using link-dependent information (e.g., relative distance or relative clock drift w.r.t. to the respective neighbor, Figure 8). In this case, the inference is facilitated thanks to the assumed reciprocity of the input data and it can be achieved without any public exchange of information.

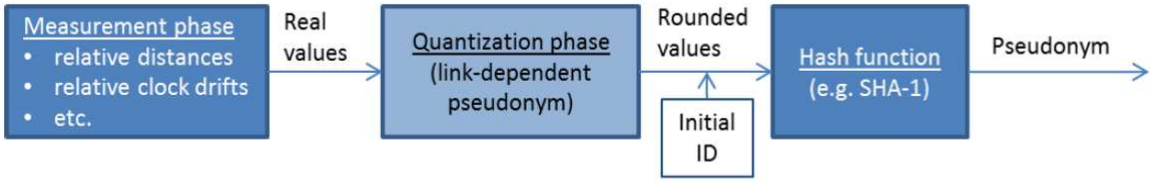

Figure 8. Typical block diagram of a LBP generator (link-dependent LBP example).

In a more practical implementation example corresponding to link-dependent LBP generation, the pseudonym (generated at A) that should be used by node B to address node A is of the form:

$$PS^{A}(A) = hash([ID_A||qd_{AB}(\Delta_d)]) \quad (2)$$

where || is the concatenation function, ID_A the initial public ID of A, Δd the distance quantization step, qd_{AB} the quantized relative distance between A and B measured at A using either n-way ranging protocols and Time of Arrival (ToA) estimation to issue RT-ToF measurements (e.g., in IR-UWB) or RSSI-based ranging using a prior path loss model (e.g., in IEEE 802.15.4).

Adding hardware device-dependent information (only for RT-ToF estimations) under the same notations as before, with $q\gamma_{AB}$ the quantized relative clock drift of A's clock with respect to B's clock (by definition, only measured at node A using the same ranging procedures) and the $\Delta\gamma$ the related quantization step, then the new link-dependent LBP is:

$$PS^{A}(A) = hash([ID_A||qd_{AB}(\Delta_d)||q\gamma_{AB}(\Delta_\gamma)]) \quad (3)$$

Inferring A's pseudonym at B without public information exchange is made possible because node B also estimates similar metrics (hopefully reciprocal) through active ranging transactions with A:

$$PS^{B}(A) = hash([ID_A||qd_{BA}(\Delta_d)||q(1/\gamma_{BA})(\Delta_\gamma)]) \quad (4)$$

An attacker E (passively) eavesdropping or (actively) exchanging packets with nodes A and B can estimate neither the relative distance between A and B (because of its different location), nor the relative clock drift, which is inherent to both the link and the hardware characteristics of the legitimate peers. Moreover, if E compromises one node, it cannot find the pseudonyms that are used on the other links. Therefore, the attacker has to make a blind (or eventually statistics-assisted) guess on the generated pseudonym.

Section 6 - Indoor localization enablers based on magnetic field data

6.1 Improvements to the nominal embodiment

As is described in the deliverable D3.1.1 [5], in our approach we consider the vectorial character of magnetic fields, and propose to use the three components (x, y, z) of a magnetic field to provide a more complete characterization of buildings.

During a first stage of our localization system, we generate maps containing the magnetic field profile of the building where the localization problem needs to be solved. During this first stage (off-line phase), orientation and user phone wearing positions are pre-established to be later considered and associated to the appropriate magnetic field maps generated at the end of the off-line training phase of our system. In this way, we generate descriptive models based on these data for user localization.

Considering each one of the pre-established orientations and user phone wearing positions, ``snapshots`` of the magnetic field are collected over short periods of time (less than a minute in each location) and throughout the building space. Such measurements are associated to the physical positions where they were gathered. Several data collection processes are carried out, considering different context conditions such as different levels of occupancy, different moments of the day, etc. Thus, the building models generated will be sufficiently representative to cover different contextual conditions.

Some modifications have been taken from the first version of the mechanism implemented. Such changes affect the phases carried out to generate the models containing the magnetic field distribution associated to the building space i.e., to the off-line phase of the mechanism. Specifically, the data clustering and classification have been eliminated of the mechanism. This is translated into an improvement in the accuracy associated to the location estimations.

6.2 Architecture and implementation updates

In Figure 9 we've represented the main modifications performed in the localization mechanism. After our first proposal of mechanism, we evaluated different algorithms and configurations, and the best results achieved in term of accuracy in localization were such associated to remove the phases of data clustering and classification. This is due to the changeable behaviour of magnetic field, which is easily affected by the presence of objects. In this sense, when no data clustering is performed, all values of the dataset are used to build the estimator, avoiding use only the data associated to certain cluster, and increasing the probability to select the closest magnetic field values to estimate new locations when variability in the magnetic field occurs.

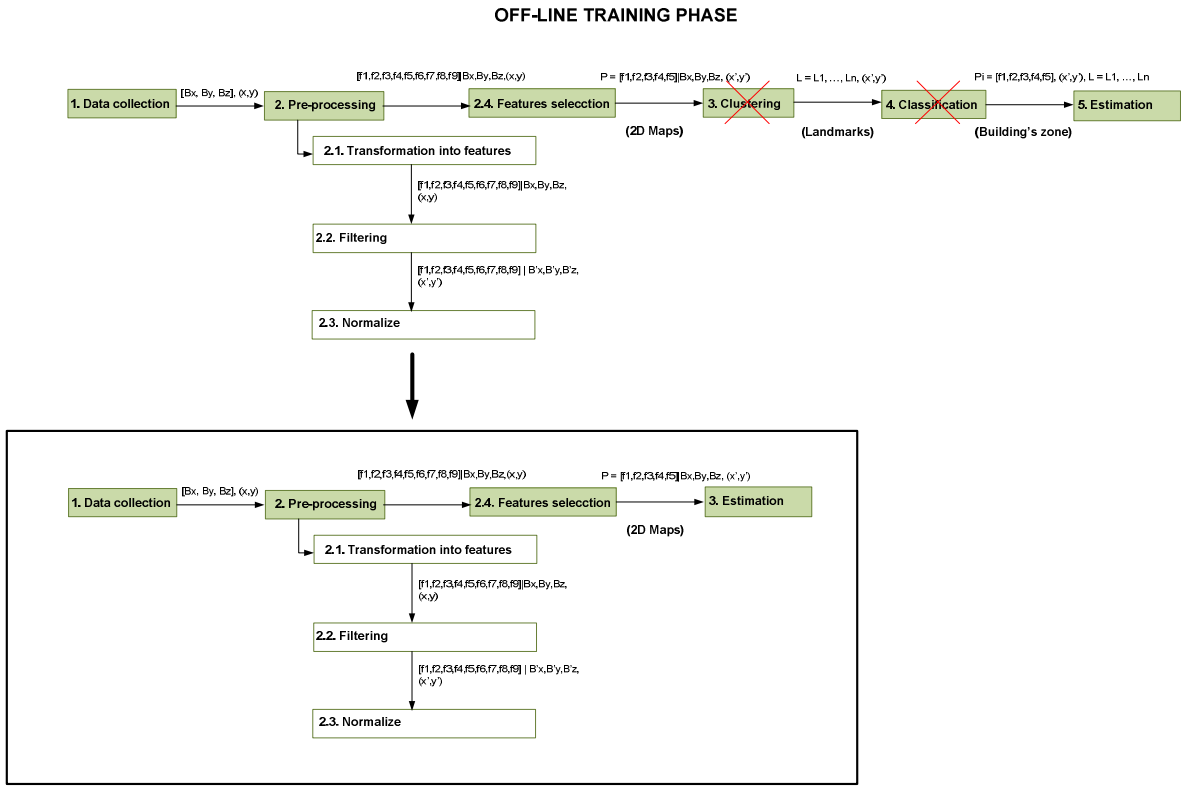

Figure 9. Improved indoor localization mechanism

Taking into account the new phases of our proposed localization mechanism, below we describe each one of them.

1. Data collection: magnetic field data are gathered using a smartphone with integrated magnetometer. It is important to take into account that such magnetometer readouts are prone to disorientation of the sensing module and are dependent on orientation and position of the user's phone during data collection e.g., the device may be in a trouser pocket or in a bag. Thus, it is necessary for the phone orientation and user wearing position to be correctly associated with the data collection performed, which will be considered in the off-line phase of our mechanism.

2. Pre-processing: this phase is responsible for preparing the measured data by transformation. Besides, feature vectors are extracted from the data for use in location estimation. Based on the raw dataset collected by the phone, during a transformation phase, compact representations of the magnetic field values, namely features, are extracted, which will be used later for localization estimation. The initial selection of features to represent magnetic field distribution is based on studies proposed in the literature that already use them for a similar purpose. The values within the dataset are grouped into windows of 64 samples, and each window is processed by several feature extraction methods, producing a feature vector that can be used to generate the clusters and train the classifier. At this stage, 27 features were extracted for evaluation (9 features for each magnetic field component: Bx, By and Bz). Then, a filter is applied to remove features extracted from the training data set that do not vary at all or that vary too much. Finally, these values are normalized in the given dataset. The resulting values are in the [0, 1] interval for every feature extracted from the initial dataset.

3. Features selection: we apply the technique of Principal Components Analysis (PCA). PCA is a widely used technique for reducing dimensionality in high-dimensional data, identifying the directions in which the observations most vary. Dimensionality reduction is accomplished by choosing a sufficient number of vectors to account for a given percentage of the variance in the original data (by default 0.95). With the aim of reducing the final computational load of the localization system, we searched the optimum number of attributes to represent the magnetic field profiles of buildings. At this point, and based on these magnetic field features, we generate maps of the building.

4. Estimation: using the knowledge available, the last step consists of carrying out a position estimation. For this, a Radial Basis Functions (RBF) network [53] is computed as regression technique, which uses all training data to estimate the user position according to its magnetic field feature vector. RBF networks find approximation solutions in the form of weighted sums of basic functions based on reference data. The main advantages of using RBF to solve our estimation problem are its scalability and easy deployment under different context conditions; where, a variable number of centroids has been identified previously.

Taking into account these phases of our indoor localization mechanism, the new schema for the on-line stage is shown in Figure 10.

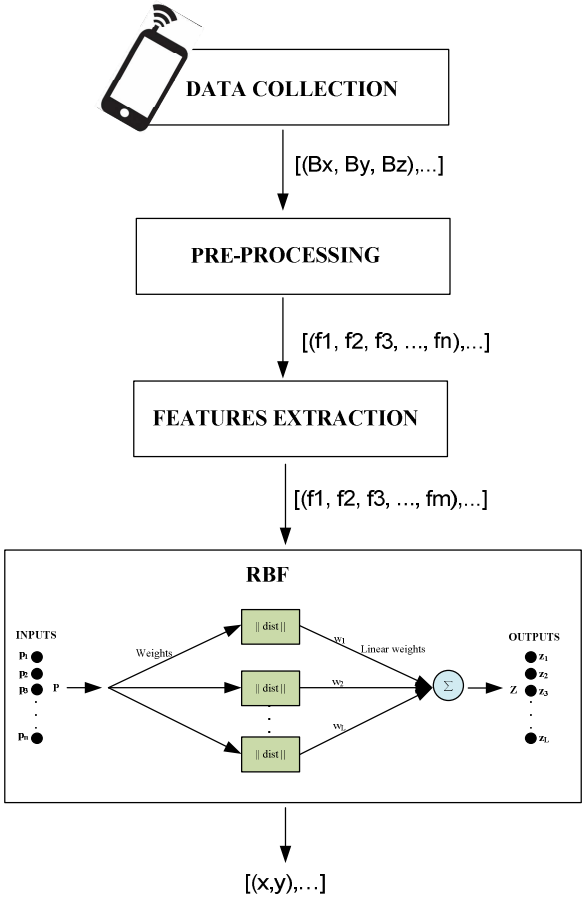

Figure 10. Building model based on the magnetic field for indoor localization

Section 7 - Final evaluations

In this section, independent evaluations are accounted still at the enabler level, keeping in mind that some of the proposals will be implemented within the overall SocIoTal framework and further validated in the frame of demonstrations (See Sect. 8).

7.1 Face-to-face enablers

This subsection provides an evaluation of our proposed approach against other state-of-the-art techniques. We first present the evaluation methodology, describe implementation details of benchmarking techniques and then present evaluation results.

7.1.1 Scenario and evaluation methodology description

The initial step for starting benchmark of DARSIS was to apply some commonly used evaluation techniques on data used to train interaction zone and proximity detection classifiers. Each technique from state-of-the-art was implemented and evaluated on the same dataset that we presented in [5] to benchmark them in a typical office environment and against a large number of samples; eight different distances 0.5m, 1m, … 4m and for each distance in three different orientations. For the classification-based approaches 10-fold cross-validation was utilised in order to determine their accuracy. To have a fair comparison all the techniques are fed with a 6-sample window.

Among the state-of-the-art techniques we have replicated Comm2Sense [40] as machine learning based solution and BlueEye [39], Free Space PLM, Office PLM [38] as examples of path loss based techniques. It should be noted that the study in [44] and Comm2Sense [40] leverage similar proximity detection technique, so the evaluation refers to both works. For PLMs given equation:

$$RSSI = RSSI_{Ref} - 10 * n * log_{10}(d) + X_{\sigma} \quad (5)$$

We computed for the specific environment RSSI_Ref = -56.7977 at 1m distance, for Free Space n_free = 2 and for Office n_office = 3.134; we also considered X_sigma = 0 for Line of Sight (LoS). Figure 2 illustrates the estimated signal strength from PLMs in comparison with empirical measurements. As shown, PLMs are estimating the correct distance, only around mean RSSI of the corresponding distance. This fact indicates that PLMs cannot follow the variation of Bluetooth RSSI measurements in real-world environments. In particular, it can be observed that BlueEye [39] follows almost the same plot with Free Space PLM. Through evaluation of proximity detection, we realised that BlueEye typically fails to resolve the complicated equation, and when solutions were found the accuracy was still very low. For this reason hereafter BlueEye method is excluded from our evaluations.

7.1.2 Considered benchmark and performance metrics

The evaluation of F2F enabler is focused on the accuracy of the approach with respect to the target classes. In this sense the performance is measured through confusion matrices, Receiver Operating Characteristic (ROC) curves and overall accuracy. Confusion matrices constitute a metric regarding the misclassification among the target classes. Furthermore, an ROC curve demonstrates the trade-off between true positives (sensitivity) and false positives (1 - specificity) of a given classifier and a given class (i.e., in our study an interaction zone or proximity). ROC curves are particularly useful when cost sensitive classifications are considered. The diagonal (y=x) line in a ROC plots represent a random selection of the classes. An ideal classifier would appear at the upper left corner of the ROC with 100% sensitivity and specificity. While confusion tables such as Table 1 provide an overview of

techniques' average accuracy, ROC curves in Figure 10 clarify the cost (i.e., rate of false positives) that each classifier imposes at a certain rate of correct classification.

7.1.3 Results interpretation

Table 1. Confusion Matrices for evaluation of interaction zone detection against state-of-the-art in percentages (%)

	DHC			DSC			Comm2Sense		
	Public	Social	Personal	Public	Social	Personal	Public	Social	Personal
Public	**88.05**	10.23	1.72	**72.88**	24.63	2.49	**35.58**	60.12	4.30
Social	4.93	**94.13**	0.94	9.16	**89.38**	1.46	5.36	**92.31**	2.33
Personal	1.16	1.10	**94.74**	2.03	2.55	**95.42**	5.02	14.76	**80.22**
	Office PLM			Free Space PLM					
	Public	Social	Personal	Public	Social	Personal			
Public	**24.15**	34.50	41.35	**40.12**	34.52	25.34			
Social	15.00	**51.80**	33.20	45.95	**37.04**	17.00			
Personal	0.02	0.23	**99.75**	0.03	12.34	**87.63**			

Tables 1 and 2 outline confusion matrices for interaction zones and proximity detection. Our DHC, DSC, DPC achieve the best accuracy with a small misclassification error in public zone and not proximity targets. Comm2Sense [40] is able to accurately detect social and personal zone, yet there is a very high misclassification percentage for public zone. Both Office and Free Space PLMs detect personal zone with high accuracy but show considerable confusion errors when resolving social-public zones and proximity or not. This misclassification error originates from reflections, causing RSSI values to overlap at certain distances (See Figure 11). As shown, machine-learning techniques are able to cope with this confusion through time dependence, as opposed to PLMs. Our DHC leveraging the hierarchical design, informative features and robust classifier overcomes the confusion between public and social zone. Environmental factors affect less in close distances i.e., personal zone, providing the ability to differentiate between personal and other zones; for that reason all approaches achieved high accuracy.

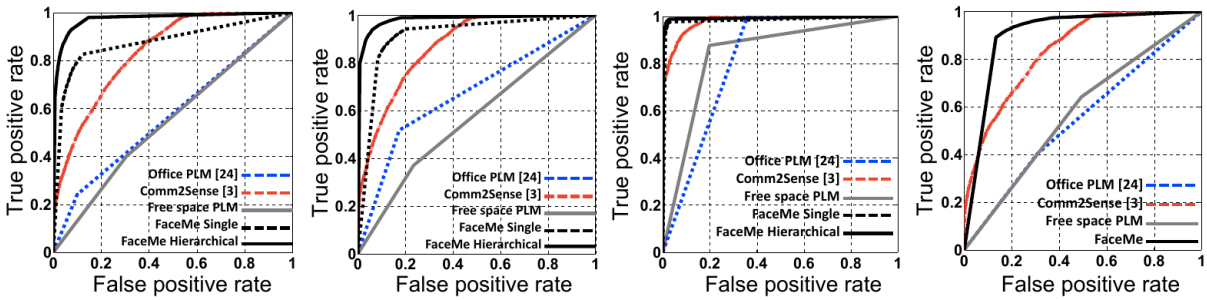

Figure 11. Comparison of F2F Proximity Detection through ROC Diagrams for state-of-the-art approaches

Figure 10 represents ROC curves of state-of-the-art techniques alongside with our proposed solutions as through discriminating interaction zones and detecting proximity from training data. Results indicate that our DHC performs close to 100% of true positives with less than 20% false positives for all interaction zones while our DPC reaches over 90% true positives with 20% false positives for proximity, showing again the effectiveness of informative features combined with a robust classifier. Comm2Sense shows a relatively high rate of false positives from the starting point

for interaction zones and proximity. However, by almost 50% of false positives, it marginally achieves perfect true positive rate in all four figures. Finally, Office and Free Space PLMs demonstrate the worst performance close to random classification in Figures 10a 10b 10d, except 10c where all the techniques achieved relatively good results because in short distances the fading effect is negligible.

Table 2. Confusion Matrices for evaluation for Proximity Detection against state-of-the-art in percentages (%)

	DPC		Comm2Sense		Office PLM		Free Space PLM	
	Not Proximity	Proximity	Not Proximity	Proximity	Not Proximity	Proximity	Not Proximity	Proximity
Not Proximity	**77.36**	22.64	**49.94**	52.06	**40.13**	59.87	**64.13**	35.87
Proximity	5.94	**94.06**	8.78	**91.22**	30.64	**69.36**	49.33	**50.67**

Table 3 represents the overall accuracy of each approach in interaction zone and proximity detection. Our approaches, DHC, DSC and DPC outperform all state-of-the-art methods. This is justified on the fact of utilising more informative and consistent features combined with a very powerful classifier. As demonstrated in this table, Comm2Sense [40] is able to attain a good accuracy with only two features for both targets. Both PLMs achieve low accuracy for interaction zone and proximity detection because of the generality of PLMs. Our DHC achieves higher accuracy, which confirms the effectiveness of the proposed hierarchical structure. Finally, DHC managed an improvement margin of at least 8% in comparison to every state-of-the-art technique.

Table 3. Overall accuracy for Interaction zone and Proximity detection in percentages (%)

	DHC	**DSC**	**DPC**	Comm2Sense	Office PLM	Free Space PLM
Zones	**93.52**	**86.76**	**-**	75.10	56.87	50.46
Proximity	**88.50**	**89.88**	**88.50**	80.39	62.05	54.03

7.2 Radiolocation enablers

7.2.1 Scenario and evaluation methodology description

Simulation-based evaluations

In this subsection, we illustrate the performance of the radiolocation enabler in case of single-link LBP generation, assessing the success rates of LBP guess at both neighbouring legitimate devices and attackers (and accordingly, the success rate of impersonation attacks based on an illegitimately guessed LBP).

Results have been obtained in three main scenarios through Monte Carlo simulations, considering 1000 distinct network realizations of 10 nodes each, with an average node degree per node between 7 and 8 neighbours and with uniformly distributed coordinates drawn in a 20mx20m area.

- First of all, we apply distinct quantization grids onto the range measurements produced by two different radiolocation technologies, namely Narrow-band (NB) at 2.4 GHz considering IEEE 802.15.4 and IR-UWB. Relative P2P distances are computed from RSSI readings and RT-ToF estimates, respectively in the NB and IR-UWB cases.
- Secondly, we evaluate the performance of pseudonym generation schemes jointly based on P2P distance and relative clock drift measurements, based on IR-UWB only.

- Finally, we evaluate impersonation detection methods based on single-link LBP using uniquely P2P NB RSSI measurements as inputs, in comparison with other conventional detection methods using the same kind/amount of input information.

The attacker model is two-fold:

- Brute-force (BF) attacks (i.e., the attacker makes a uniform random guess on the value of the internode distance and/or relative clock drift);
- Probabilistic brute-force (pBF) or statistics-aided attacks (i.e., the attacker knows a priori the distribution of the true internode distances in the area and makes his best guess accordingly).

In our simulations, we consider assigning to all the feasible links some random radio channel configurations (i.e., Line of Sight (LOS), Non Line of Sight (NLoS) and severe NLoS (NLoS2)), along with their corresponding radio parameters (i.e., ToA standard deviation for RT-ToF vs. shadowing standard deviation, reference path loss and path loss exponent for RSSI), depending on the actual internode distance like in [57]. Once each link configuration has been allocated, one can use the associated conditional model parameters to estimate the range. For NB RSSI-based ranging, we consider the median estimator from [16], with a path loss exponent α = [1.7; 3; 5] and as a shadowing standard deviation σ_S = [0.5; 3; 5] dB, respectively in [LoS; NLoS; NLoS2] channel conditions. For IR-UWB RT-ToF estimates, we consider the standard deviation of ToA estimation as [1; 2; 3] ns (i.e., characterizing the dispersion of the noise terms affecting the TOA estimates for each received packet involved in the n-way ranging transaction), according to empirical observations from [58]. Relative clock imprecisions are assumed to be bounded by $\pm\delta$ = 20 ppm (worst case), which is representative for low-cost embedded oscillators.

Measurement-based evaluations

Complementing our evaluations, we have also exploited real data from a measurement campaign initially devoted to wireless localization. The latter campaign involved real devices enabled with the IR-UWB technology (among others) [59]. Thus one aim was to get proof-of-concept validations to the proposed location-based device's authentication and identity reinforcement strategies (i.e., out of real radio signals and in a representative indoor wireless environment).

The involved IR-UWB devices allow RT-TOF measurements between peer-to-peer devices [58] over communication links at around 347 kbps. They operate at the centre frequency of 4.5 GHz over a bandwidth of 500 MHz, in compliance with the European spectrum regulation in the so-called "lower band" [3:5]GHz. These nodes provide ranging timers associated with intermediary Time of Arrival (ToA) estimates within the time resolution of 1 ns, which can be directly converted into RT-TOF, and hence into a relative distance. On the medium access control layer level, specific beacon-enabled TDMA ranging transactions.

Up to 16 of these IR-UWB devices were deployed in a mesh topology in a 30mx18m typical indoor environment (professional office environment), mixing LOS and NLOS radio obstruction regimes. In this scenario, the average connectivity degree per device (i.e., number of available 1-hop neighbours) was observed to be on the order of 7-8 nodes but theoretically up to 15. For each addressable 1-hop radio link, tens of successive bilateral/reciprocal acquisitions have been performed between pairs of devices, enabling to test both link-dependent LBP generation (i.e., 1 LBP out of 1 relative distance w.r.t to 1 neighbour) and position-dependent LBP generation (i.e., 1 LBP out of the 2D coordinates of the device and the set of relative distances w.r.t all its 1-hop neighbours). In the position-dependent LBP case, 2D coordinates could be determined through non-cooperative trilateration positioning, using at least 3 anchors among the available neighbours.

Similarly to simulation-based evaluations, we have evaluated:

- The probability of success per link for the legitimate guess in the link-dependent LBP case, averaging over more than 60.000 realizations (i.e., over different devices and neighbors, over all the occupied positions, and over all the acquisitions for a given pair-wise link);
- The probability of success of Brute-force (BF) attacks (i.e., the attacker makes a uniform random guess about a unique quantized value of the inter-node distance in single-link LBP or uniform guesses jointly about the 2D coordinates and an (a priori unknown) number of relative distances with respect to the 1-hop neighbors of the legitimate node under attack, using the same uniform quantization grids for both range and position information);

7.2.2 Considered benchmark and performance metrics

First of all, for both simulation-based and measurement-based evaluations, in order to assess the robustness of our proposal with respect to noise over legitimate links (from the legitimate guess perspective for link-dependent LBP), as well as its resistance against attackers (from the illegitimate guess perspective for both link-dependent and position-dependent LBP), we define two different performance indicators:

- The probability of a successful inference of the link-dependent pseudonym by a legitimate user based on its own relative ranging observations:

$$P_l = \Pr\left[PS^B(A) = PS^A(A)\right] \quad (6)$$

- The probability of a successful guess of the LBP by an attacker (evaluated for both BF or pBF with link-dependent LBP and BF only with position-dependent LBP):

$$P_{(p)BF} = \Pr\left[PS^E(A) = PS^A(A)\right] \quad (7)$$

In simulation-based evaluations, so as to assess the performance directly in terms of impersonation detection, we also assess the probability of a successful attack as the probability miss-detecting an attacker E (pretending to be A or B), as a function of its relative distance with respect to the legitimate node A under attack (i.e., the distance dAE), while averaging over all the possible distances between the legitimate nodes A and B (dAB). We compare the proposed RSSI-based single-link LBP approach (i.e., the probability that the attacker claims the rights LBP without being detected) with a state-of-the-art RSSI Similarity based Authentication (SA) method [18], which aims at detecting large unexpected RSSI transitions (over time) between consecutive frames at one legitimate receiver from another legitimate transmitter (i.e., due to the insertion of impersonated packets from the attacker). We assume that the SA threshold for impersonation detection at the legitimate nodes is set at $\pm 3 \times \sigma_S$ around the true (presumably known) mean RSSI value.

Additional (analytical) mathematical details about the derivation of the attacker success, in terms of both illegitimate guess (LBP inference) and impersonation attack miss-detection, can be found in [67].

7.2.3 Results interpretation

Simulation-based results

On Figure 12, we represent the legitimate agreement probabilities for distance quantization from NB and IR-UWB estimates with two radio link conditions (only LoS and a realistic mixture of LOS, NLoS and NLoS2). Pseudonym generation from IR-UWB RT-ToF estimates is thus shown to be more robust

to measurement noise dispersion, regardless of the channel conditions. We also report the brute-force and the probabilistic brute-force successful attack probabilities as a function of the distance quantization step and identify advantageous quantization steps (defining an operating domain) that maximize the gap between the legitimate agreement and the successful attack (i.e., illegitimate guess) curves (e.g., Δ_d = [2… 10] m).

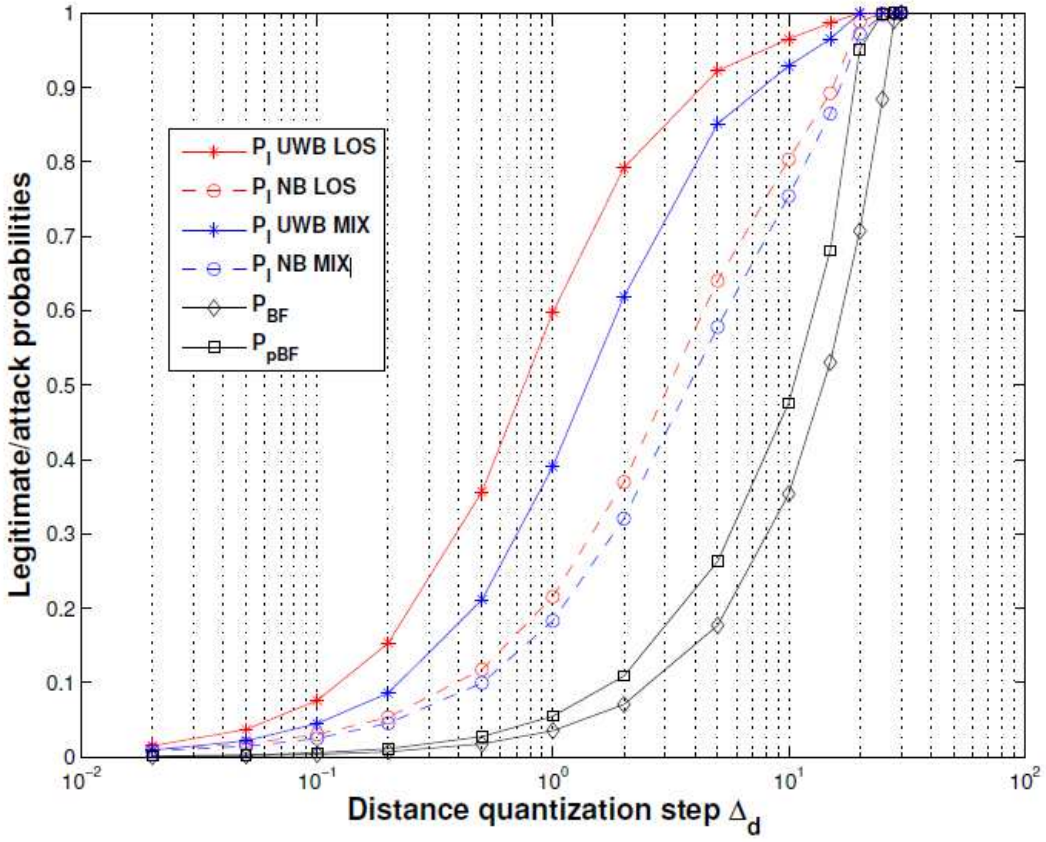

Figure 12. Performance of link-dependent range-based pseudonym generation: legitimate inference success probability (P_I) and attack success probability ($P_{(p)BF}$) as a function of the range quantization step Δ_d.

After incorporating the relative clock drift into the quantization procedure (e.g., with Δ_d = 10 m for the sake of illustration) with IR-UWB radio only, both the legitimate agreement and the successful attack probabilities tend to decrease, as shown on Figure 13. Nevertheless, incorporating the drift is always beneficial with respect to the probabilistic brute-force strategy as it can be seen at relatively high values of the clock drift quantization step. The success rate of the legitimate guess with relative clock drift becomes close from that obtained with no clock drift information, while the successful brute-force attack probability (performing brute-force guess about the measured relative clock drift on top of the statistics-aided guess about the range measurement) is always lower than that of the former statistics-aided attacks when no drift information was used. So overall, incorporating drift information shows better immunity against attackers with a priori statistical knowledge about the expected relative range distributions between the legitimate devices.

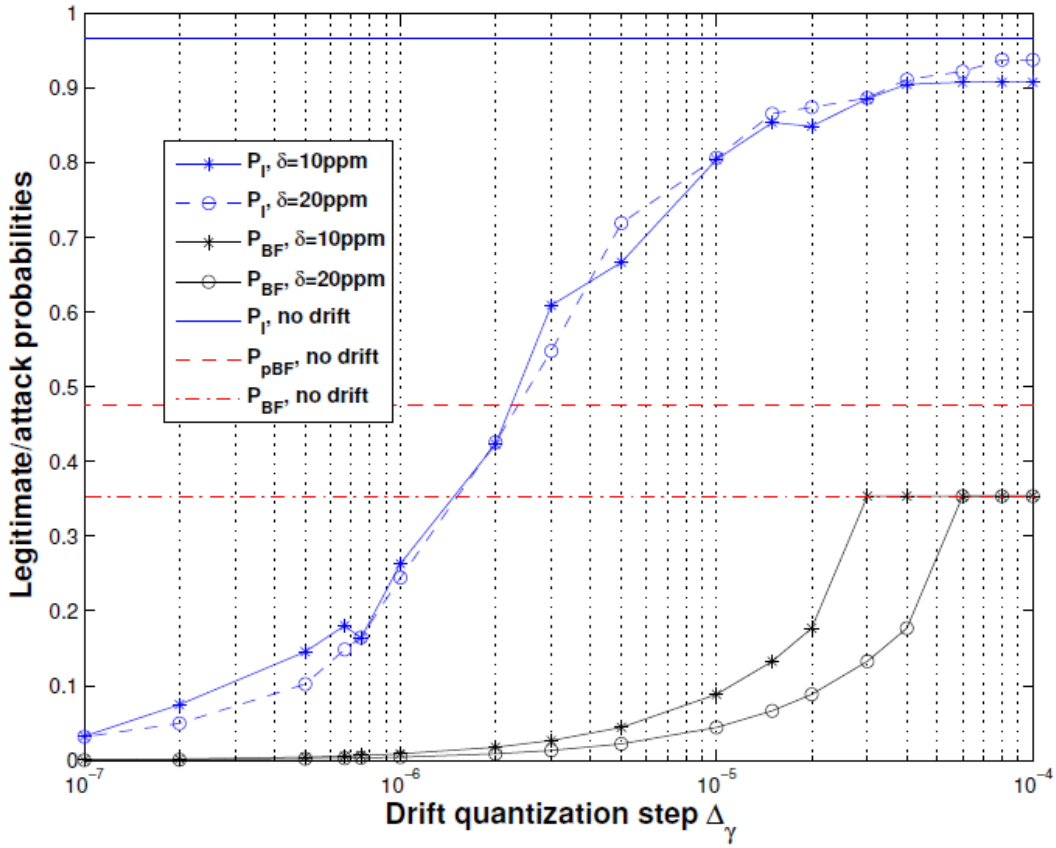

Figure 13. Performance of link-dependent LBP generation based on relative range vs. jointly relative range and clock drift measurements (Ex. with fixed range quantization step Δ_d = 10 m).

Finally, as shown on Figure 14, the impersonation attack success with the proposed distance-based pseudonym generation (PBF) is lower than that in the classical RSSI-based authentication approach (P_{SA}), for realistic medium-to-large shadowing variances (3 to 5 dB) and for practical ranges from the attacker, even with relatively large Δ_d.

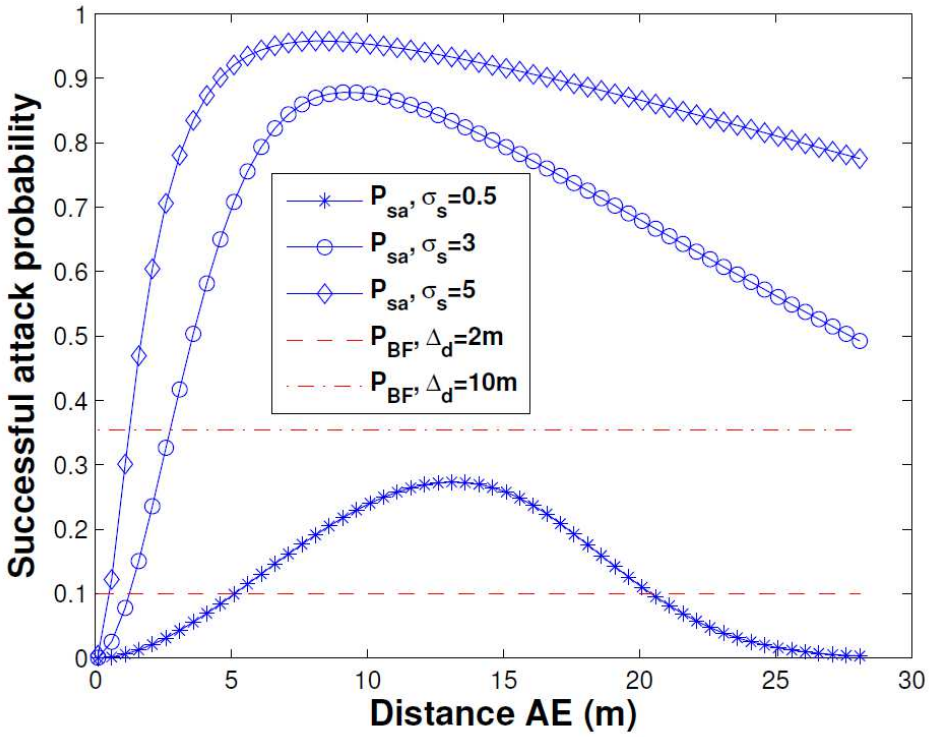

Figure 14. Comparison of the successful impersonation attack probabilities in pseudonym generation vs. direct RSSI monitoring (SA) for different attacker-legitimate distances d_{AE}.

Measurement-based results

In the first scenario, we assess the performance of link-dependent (range-based) LBP generation in terms of both legitimate inference success probability and Brute-Force attack success probability, as a function of the quantization step. On Figure 15, one can thus note that the legitimate guess success rate observed with real devices and measurements is significantly higher than that expected in the most optimistic simulation case (See Figure 13 in LOS), whereas the Brute Force performance is by definition the same as before. This likely comes from the fact that ranging errors in NLOS cases are strongly but reciprocally biased (i.e., the bias value being large, but constant and reciprocal on both sides of the link), although the standard deviation of their random dispersion is quite low and approximately the same as in LOS (i.e., instead of assuming independent on both sides of the link in our simulations). This tends to strengthen the reciprocity of the input measurements, and hence, of the legitimate LBPs.

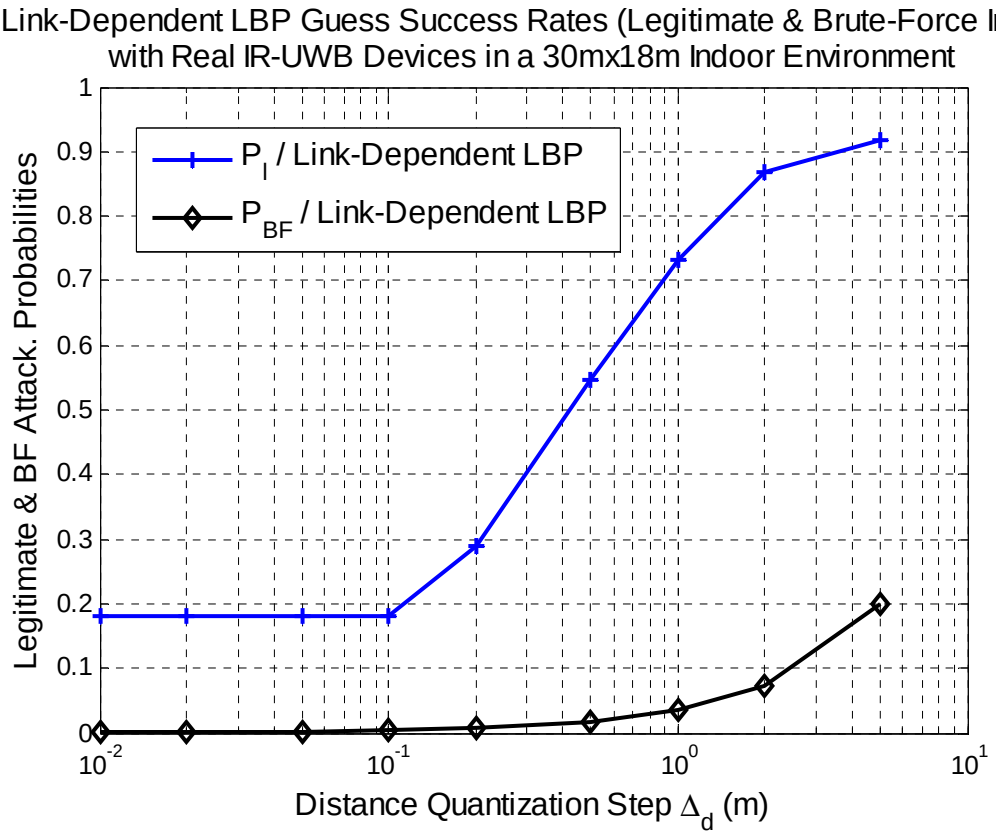

Figure 15. Performance of link-dependent range-based pseudonym generation (i.e., using 1 RT-ToF Measurement wrt. 1 Single Neighbor as Input): legitimate inference success probability (P_l) and attack success probability (P_{BF}) as a function of the range quantization step Δ_d.

In the second scenario, assessing the immunity of position-dependent LBP generation against blind illegitimate guesses from an attacker, it is shown that the inclusion of the set of relative distances with respect to an arbitrary number of neighbors (on top of 2D coordinates) when generating the LBP is much more difficult to infer from an attacker than in the link-dependent case (See e.g., Figure 16). This naturally results from the explosive combinatory complexity in lack of precise information about the actual physical connectivity of the device under attack (i.e. the exact number of available neighbors and thus, the actual number of independent range measurements incorporated as inputs while generating the LBP). For that reason, whereas the link-dependent LBP verification is mostly intended as authentication overlay (i.e., completing other conventional methods), the position-based LBP verification, requiring a centralized entity to store LBP information as part of the acquired context, could be used as primary means, provided that communication links from the devices to the centralized entity are authenticated and secured and that LBP are renewed after usage/disclosure.

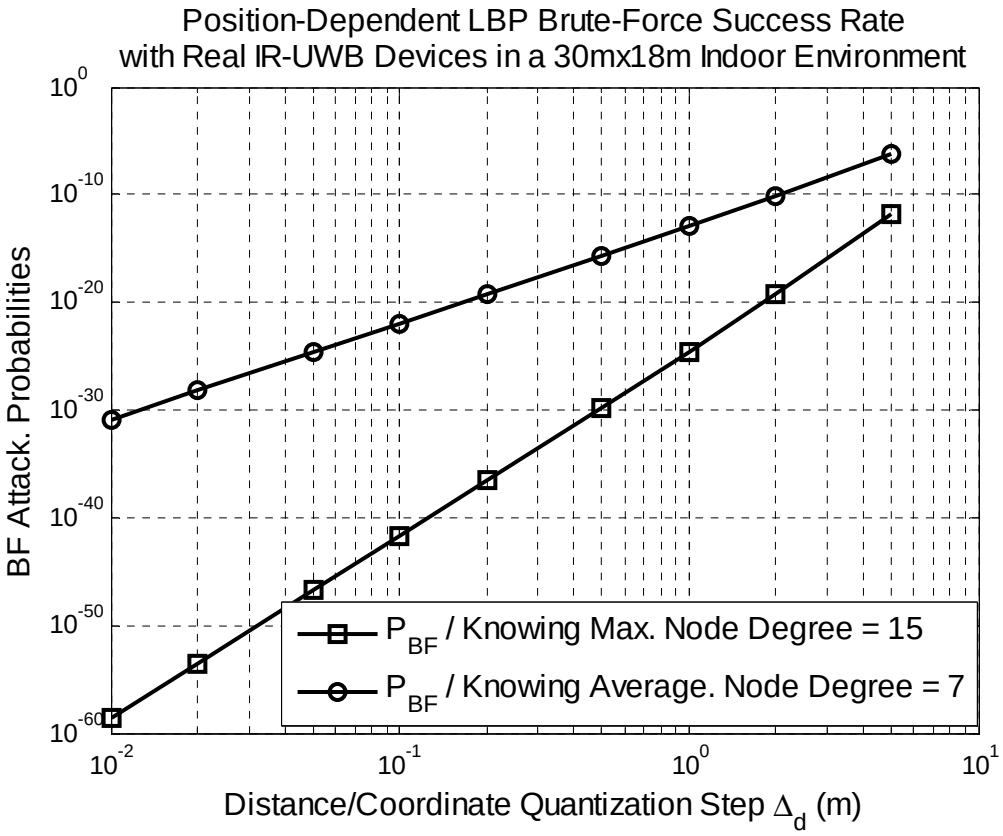

Figure 16. Immunity of position-dependent pseudonym generation (i.e., using 2D Node Coordinates & Set of RT-ToF Measurements wrt. Neighboring Nodes as Inputs): attack success probability (P_{BF}) as a function of the range/position quantization step Δ_d (assuming knowledge of the average or maximum nb of neighbors)

7.3 Magnetic localization enablers

7.3.1 Scenario and evaluation methodology description

In order to evaluate our indoor localization mechanism, firstly it is necessary to choose the optimum parameters of design of the different techniques that compound the mechanism. For this, in the target building where localization problem wants to be solved, it is necessary to collect data of the magnetic field distribution throughout the building space of interest. Using the data collected, an optimum configuration for each technique involved can be obtained after the analysis of the results associated in term of localization error. Therefore, in this section we describe the experiments carried out in the building of the Computer Science Faculty of the University of Murcia to get the optimum parameters to implement the localization mechanism proposed as well as the validation of the mechanism with 10-fold cross validation over the training dataset.

We have developed a sensing application on an Android com HTC Sense for the HTC One X (S720e). This phone is equipped with a Hall-effect geomagnetic sensor3 in three axes. The sensor implements a Dynamic Offset Estimation (DOE) algorithm to automatically compensate the magnetic offset fluctuations, thereby making it more resilient to magnetic field variations within the device [56]. In addition, the effect of high frequency ambient noise is mitigated by averaging the measurements prior to phone calibration. Our application is able to gather magnetometer signals with frequency of 25 Hz and record them into a database allocated in an external platform (in Dropbox in our case).

Ten subjects were selected from the Information and Communications Engineering Department of the University of Murcia to perform the experiments for which the data were collected. In this way, the data collected cover different user paths inside the building in a same moment of a day and in different days. In the University of Murcia there is not any ethics requirement for experiments with humans, but the experiments performed respected every aspect related with the privacy and confidentiality of the participants. During the data collection, the subjects were asked to walk on predefined trajectories along the first floor of the Computer Science Faculty. Walking along these baseline trajectories was performed while users carried their phones in their hand and in a fixed orientation with respect to a reference system of coordinates. All participants carried their phone in a same position and considering a same phone orientation.

Since data collection was performed on different days and at different moments of the same day, variability in the context conditions is included in our base data sets. Therefore, the size of the final data set considered for testing is of 1065 measurements. Then, using this data set, the data processing techniques presented in Section 6.2.2 were analyzed considering different values for their implementation in Matlab.

7.3.2 Considered benchmark and performance metrics

To provide a detailed analysis of the results achieved by our localization mechanism, we focus on the localization results obtained for the first floor of the Computer Science Faculty depicted in Figure 17. The magnetic field distribution along this floor does not present high variability compared with the zones where the lifts are located. Therefore, the location results achieved in this floor cover the cases of buildings where there is not the best contextual conditions to apply a localization solution following the approach of using the magnetic field sensed inside.

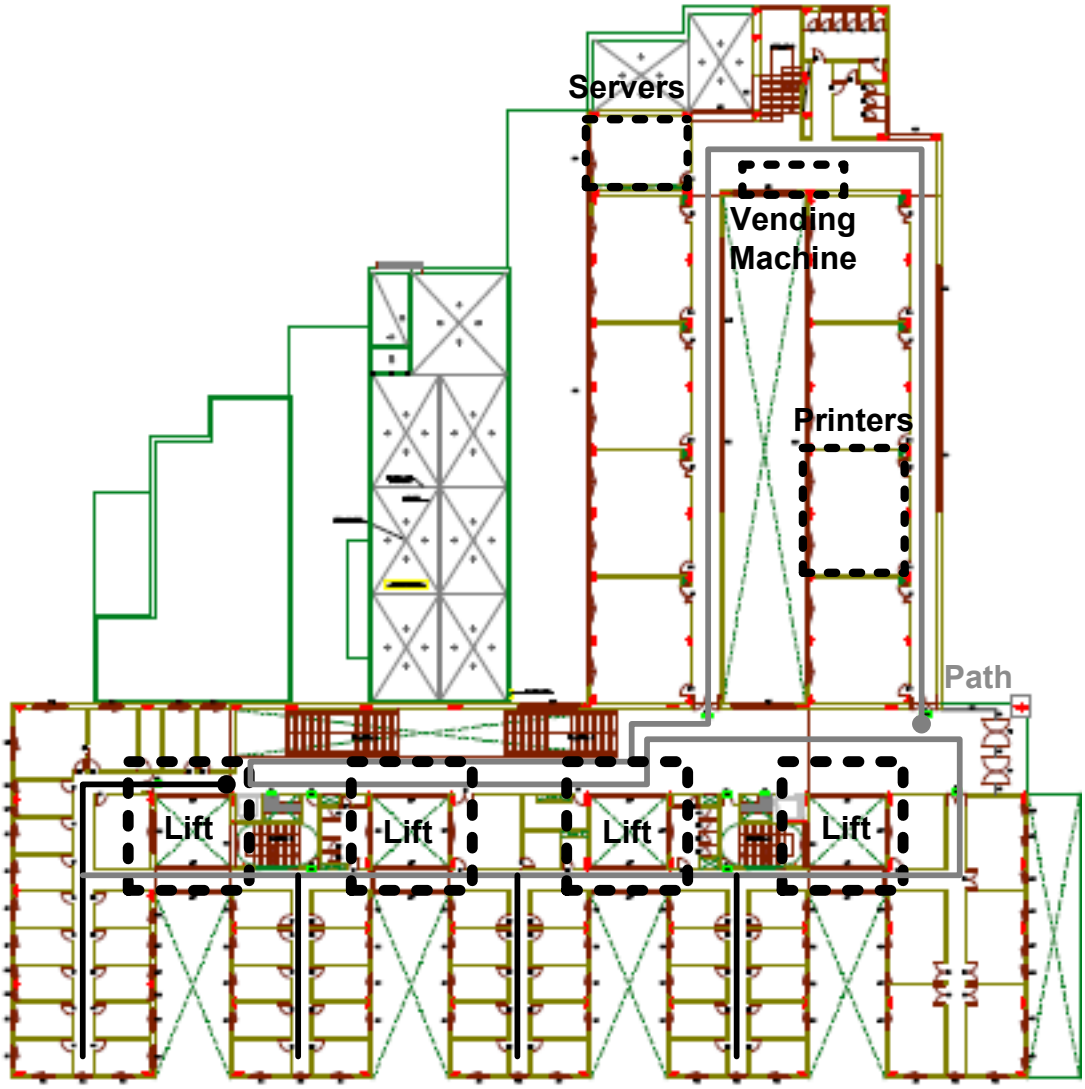

Figure 17. First Floor of the Computer Science Faculty of UMU

As results of the offline training phase of our localization system, a 2D map containing magnetic field features of the building and the RBFs were obtained. This map was associated to a predefined phone orientation and position in which participants carried their phone.

By considering the map of the building containing the magnetic field profile resulting from the training phase, the classification mechanism charged with assigning to each new measurement the zone where it belongs was evaluated. Firstly, we obtained the mean and deviation values of the accuracy achieved by each RBF implemented for estimating the user position. Table 4 shows the results.

As can be seen, it is possible to achieve very accurate localization results, with a mean value of 3.9 m and deviation of 2.7 m. However, note that the number of different sources of magnetic field perturbation in the scenario under analysis is not high, which is related with the main drawback of the solutions that follow the approach of using magnetic field measurements for indoor localization.

Table 4. Accuracy in location estimation and accuracy deviation

Accuracy in location estimation (m)	Accuracy deviation (m)
5.3	4.0
1.5	1.2
0.4	0.2
5.5	3.3
4.1	3.0
3.8	4.3
2.1	1.5
2.1	2.1

To validate the results achieved by our mechanism, we take as reference the most similar work presented in the literature [54]. In this work, the Nearest Neighbour (NN) algorithm [55] considering Manhattan distance is used for estimating user positions, i.e., as regression technique. We compare the results obtained considering such regression technique with the results associated to our proposal, in which we use RBFs as regression mechanism. As in our approach, the authors also propose using the three elements of magnetic field measurements sensed inside building to solve the indoor localization problem.

Now we focus on the localization results obtained for the corridor depicted in Figure 18. This corridor presents a high human activity level due to the numerous laboratories allocated there. Besides, it is one of the longer zone of the floor selected for the tests (28 m).

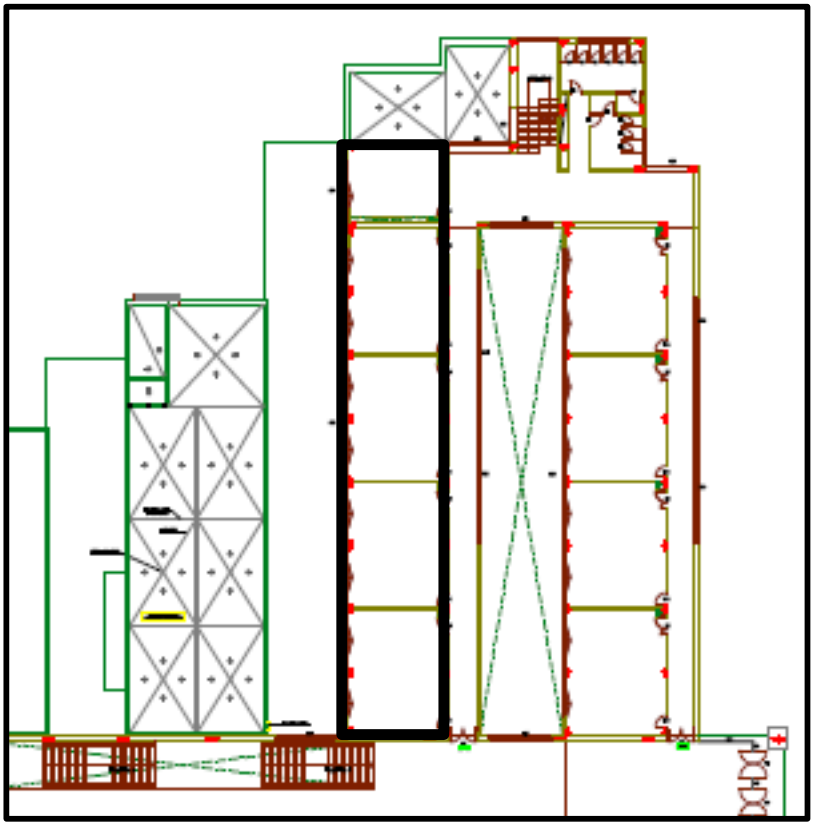

Figure 18. Magnetic field landmarks identified in a corridor

Based on our data set of the magnetic field sensed in this corridor, we apply the NN technique to estimate user positions once the building zone where user is located is known. We present the results of this comparison in terms of accuracy. As result, the RBF achieves a mean accuracy of 3.9 m, whereas a value of 5.7 m is achieved by NN. Therefore, our approach applying RBFs to solve localization estimates improves on the results provided by related studies.

Finally, for a more complete assessment of the proposal to use the magnetic field for indoor localization, we present a comparison of the localization results achieved by our mechanism with those provided using another phone-based technology and considering the same test scenario (i.e., the corridor shown in Figure 18). With this comparison we intend to validate the results obtained with our proposal of indoor localization system comparing with the results of another system already validated as feasible solution in indoor environment and currently being used for solving such problem.

Following the approach of using WiFi signals for indoor localization, T. Garcia-Valverde et al. proposed the localization system presented in [46]. This proposes a localization system which receives WiFi signals from a number of existing WiFi access points with no prior knowledge of the location of the access points and the environment. This system provides the percentages of success in the classification performed to predict location. Therefore, its level of granularity is in term of building's zones.

Note that the WiFi-based localization system was developed in the University of Murcia, which was evaluated in different buildings of the University and providing the most accurate results in the building of the Computer Science Faculty, since in this building the number of access points deployed and available to be used by this system is high.

In Table 5, we show the classification success results obtained with WiFi systems and our version of the mechanism including clustering and classification.

By considering the surface of the target zones, we can provide an approximate error for the WiFi-based localization system. The mean values are 73% for WiFi and 75% for magnetic field measurements. Considering the surface of each zone involved in the classification, a mean error of 7.6 m can be obtained using WiFi, and 6.1 m using magnetic field. But note that after this classification, our localization system provides more accurate estimates applying a regression technique to the magnetic field data associated to the zone resulting from the classification.

Table 5. Success in location classification considering WiFi and Magnetic Field

Accuracy in location estimation	Success using MF	Success using WiFi
6.1m x 6.1m	88%	75%
3.2m x 3.2m	87%	54%
3.2m x 3.2m	70%	57%
3.2m x 3.2m	65%	76%
3.2m x 3.2m	68%	92%
3.2m x 3.2m	73%	88%
3.2m x 3.2m	75%	68%
3.2m x 3.2m	74%	73%

7.3.3 Results interpretation

Analyzing the results obtained from these two phone-based solutions to indoor localization, it can be seen that the WiFi-based system is more sensitive to the problem of adjacent zones, most of the classification errors occurring in positions close, but belonging to different adjacent areas. This is mainly due to the variable distribution of wireless signals in an indoor environment. However, the magnetic field-based localization is more sensitive to resolve the problem of distinguishing zones where the magnetic field variability is low, as happens between zones 3, 4 and 5 in Table 5.

Based on the assessments of the proposed indoor localization system, we conclude that magnetic field measured by the magnetometers integrated in smartphones represents a feasible and accurate solution for solving localization problems in buildings containing perturbation sources of the Earth's magnetic field, such as lifts, electronic devices, machines, etc. Of note is the fact that the reference building used for testing presents a medium level of magnetic field perturbation, so that the results provided in this context are applicable to similar types of building. Furthermore, the mechanism proposed to generate magnetic field profile maps, the classifier design and the estimation process are totally reproducible for any building.

7.4 Overall assessment of enablers performance

As described above, each enabler performance evaluation mainly focused on assessing the accuracy of the provided solution. For all the approaches, with respect to D3.1.1 [5] each enabler evaluation has been improved along the following dimensions.

For the F2F enabler, the evaluation of accuracy of new features, namely distance estimation, has been performed and compared against state of the art approaches. In this case, the simulation-based evaluation has been preferred with respect to real-world based scenarios in order to allow comparison with a number of different solutions in easy-to-replicate environments. This complements the previous analysis where real-world evaluation of the approach accuracy in detecting F2F relations was performed.

For the radiolocation enabler real-measurement based evaluations have been considered together with simulations, thus confirming the accuracy and resilience to security attacks of the proposed solutions in real application scenarios.

Finally, the magnetic location indoor positioning enabler has been evaluated and compared in real-world environments against existing state of the art solution, thus showing its suitability for the SocIoTal envisioned application scenarios.

As result of the performed evaluation, the designed enablers showed enough maturity and improvement with respect to previous solutions, thus guaranteeing adequate level for the quality of the information they will generate and provide to other SocIoTal components. Their integration in the SocIoTal platform, as described below, will also allow to evaluate this on the field and to also focus on different evaluation aspects, i.e., energy efficiency of the implemented solutions. As each enabler will be implemented and used in end-user devices, sometime running also on constrained resources, efficiency is a crucial aspect to address, in order not to compromise system performance and end-user experience. For the enabler included in the SocIoTal field trials such aspects will be further evaluated during WP5 activities and reported in relevant work package deliverables (D5.2 [9]).

Section 8 - Enablers usage into the SocIoTal framework

8.1 Face-to-face enabler

8.1.1 Fulfilment of SocIoTal requirements and needs

In light of its nature of representing an enabling technology, the SocIoTal F2F enabler will provide a software component implementing a dedicated IoT Service. It is envisioned that such service will be deployed on devices that are part of the SocIoTal platform. Because the information extracted by the F2F IoT Service relates to the user personal sphere, it appears clear how mobile smartphones are the target devices that will benefit of such information, when participating in the SocIoTal platform. The information provided by the F2F IoT Service will be exploited in different ways, in order to make more secure and privacy preserving the participation and information shared by SocIoTal devices, thus satisfying the envisioned platform requirements. More details are provided below.

From Face-to-face to context

As described above the F2F enabler will extract information about the nature of social relations incurring among people, by classifying them accordingly to the users orientation and interpersonal distance. Such information can be used in a number of different ways. First of all, by extracting periodically or according to well-defined events, the F2F information, a more accurate characterization of the device context can be provided. Such characterization will allow from one side to better classify the environment surrounding a given SocIoTal device (e.g., mobile smartphones) in terms of discovered surrounding devices and relations with them. As SocIoTal devices, the smartphones are seen in terms of their capability to provide additional IoT Service, related to the production and consumption of the information generated by their embedded sensors and users in supporting the creation of citizen-centric services. The extracted context will be then used from one hand to locally decide which information the device could share, in terms of available IoT services and according to the detected nearby devices and nature of discovered relations. On the other side, the extracted context information can be used to annotate every information generated, thus allowing to guarantee access to it according to specific authorization policies, globally applied by the SocIoTal Authorization component. For instance, specific information cannot be shared if the originating device is detected to be in a context in which it is surrounded only by device in an intimate relation (i.e., home context).

From context to trust

On the other side, the information about F2F relations, generated by each SocIoTal enabled smartphone devices, can be used and analysed to establish relations about different devices and their users. Such information will be fed to the SocIoTal Trust and Reputation Manager (T&RM) in order to allow computation of reputation score for classifying the recurrent social relation incurring between two given users and their devices. By relying on such information, each SocIoTal device will be able to create its user social graph, by relying on the type and frequency of real social encounters. According to the type and frequency of the relations, different reputation can be associated to different devices and evolve over time, while new information is acquired. By using the information extracted by the F2F enabler and consuming it locally, each device will be able to extract this reputation score and share it accordingly with the T&RM or the shared context information can be consumed by dedicated infrastructure module to extract and update similar knowledge. According to this, the request and sharing of information generated by a given device can be allowed to only specific other devices which comply to a defined trust and reputation score. For instance, specific information cannot be shared when asked by devices, which never previously shared an intimate social relation with the considered one (i.e., device owned by office colleagues).

A more formal presentation and integration of the F2F enabler information into the SocIoTal architecture is provided in the following subsections.

8.1.2 Related IoT service

According to the SocIoTal architecture, the F2F enabler provides IoT Services, generating and exposing a well-defined type of information, accessible through specified APIs. As described above, the Authorization and Trust & Management blocks consume this information. The enabler data available for creation, retrieval, modification and deletion are briefly summarized below (See also Figure 19).

The Face-to-face interaction detection enabler incorporates two concrete IoT Services:

- *DirectionData*. This IoT Service retrieves the facing direction of the users i.e., direction of front part of user's torsos, inferred by the walking locomotion of the user. The knowledge of user's facing direction will lead to the computation of users' relative orientation. The relative orientation is one of the key parameters of the face-to-face interaction detection.
- *NearbyDevicesData*. This IoT Service retrieves information about nearby devices of a user, by utilising a communication mean such as BT discovery that acquires data including signal strength, device details and similar. Given the data collected, the enabler is capable of estimating the interpersonal distance of the users, thus inferring the occurrence of a face-to-face interaction or not.

These two IoT Services convey data to Virtual Entity (VE) service for face-to-face interaction detection and several other components of the system.

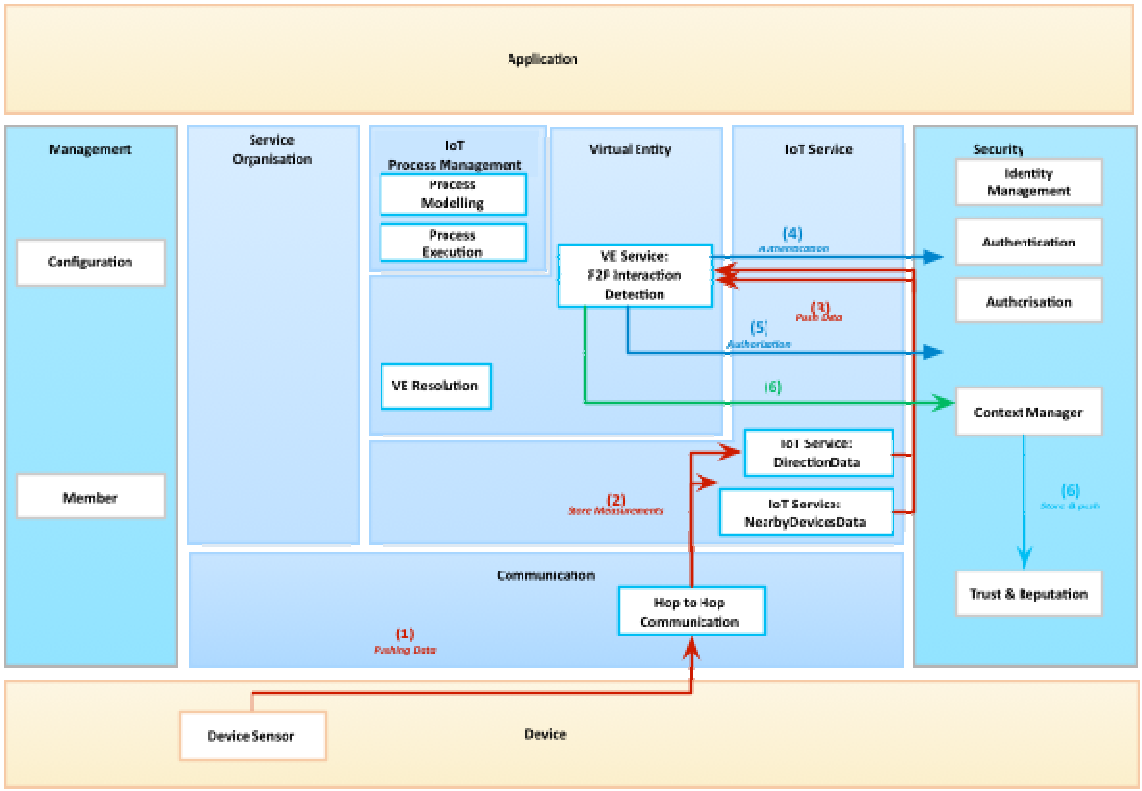

Figure 19. Face-to-face enabler architecture compliant with IoT-A (functional/information view)

The combined information provided by the VE as well as the simple atomic information provided by the implemented IoT Services can be accessed either directly (using dedicated APIs) or redistributed in the form of context information shared through the Context Manager.

8.1.3 Interfaces with authorization, trust and reputation management blocks

The F2F enabler will extract information about social relations with surrounding devices. Such information can provide contextual information to be used by the Trust Manager to verify the trustworthiness of a subject device B with respect to the considered device A. This information will be used when services (e.g., data sharing) are requested to device A. In addition this information can be used to verify rules requested by the Authorization Manager, i.e., based on device detected context (e.g., the device is in a public environment, surrounded by many untrusted (non)SocIoTal devices).

The context information that needs to be extracted (e.g., using the Context Inference module) and shared (e.g., using the Context Communication module) can be modeled as on the figure below.

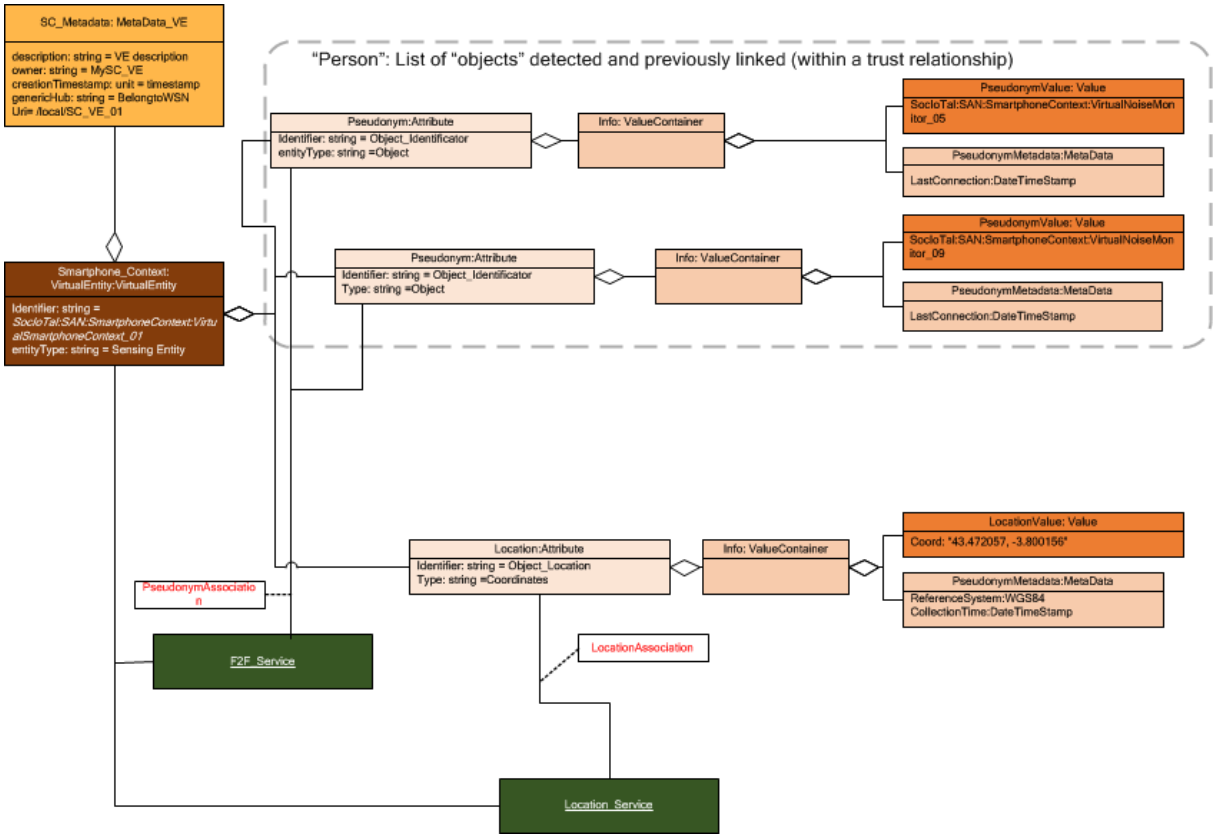

Figure 20. Context information & model for the F2F enabler

The F2F enabler is able to recognize on-going real-world interactions. Further, by inferring the interaction zone in which the social interaction is taking place, it is able to estimate the social relationship among people. Information about the entities involved in the social relation is provided using the pseudonyms associated to corresponding SocIoTal devices and annotated with information about the type of relation (e.g., personal, social and public) and location where is taking place.

In case not-SocIoTal devices are present, because a proper estimation of F2F social relations cannot be performed, a predefined pseudonyms will be assigned to the discovered devices and the relation classified as unknown, thus still providing enough information to classify the context surrounding the

considered SocIoTal device. The information provided by the F2F enabler can be seen as standalone information classifying the context of the device or as set of attribute associated to the sensing data generated by the considered device and characterizing their context.

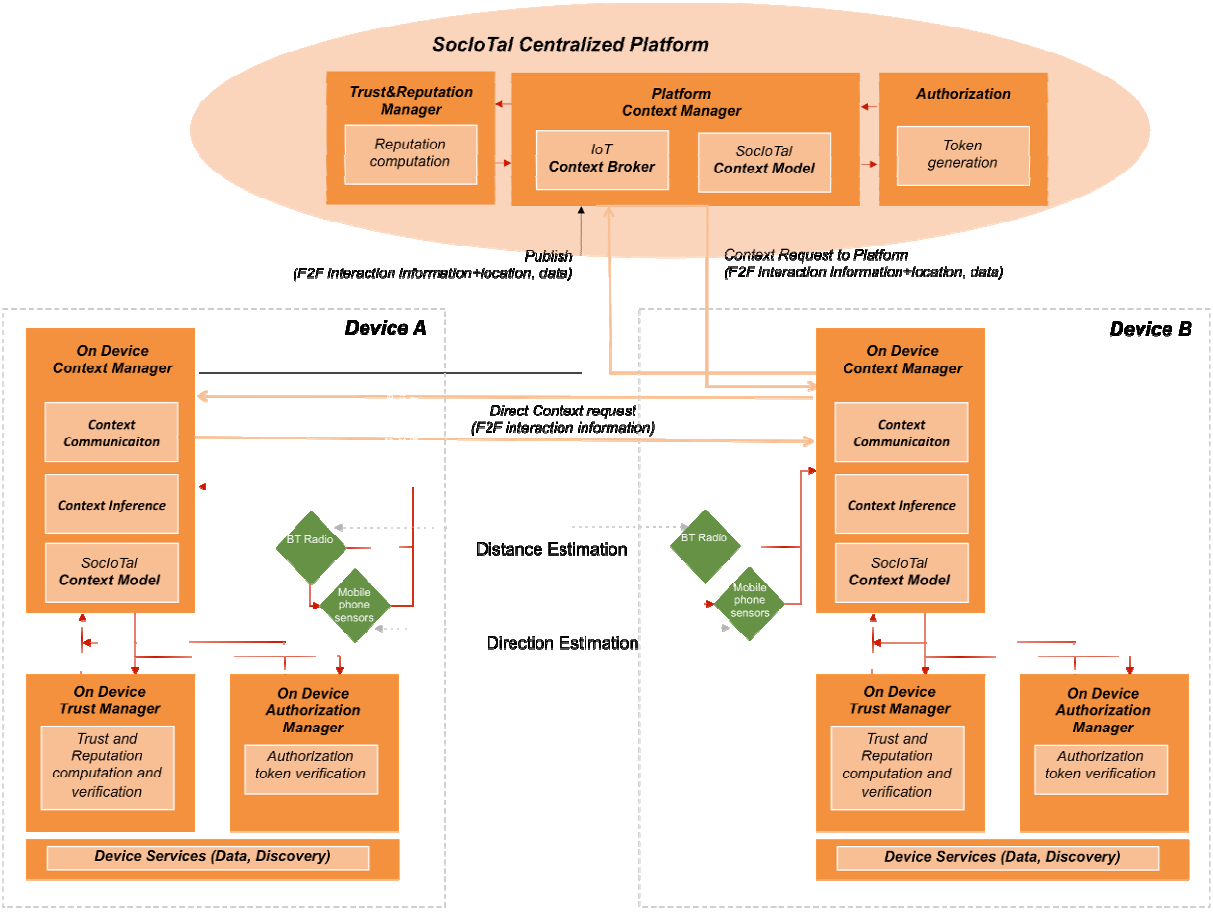

Figure 21. F2F Enabler integration

There are two ways envisioned in the SocIoTal architecture to allow access to services, such as data sharing, provided by SocIoTal devices. In both cases, privacy-preserving data access is regulated according to the type and number of detected social relations with targeting and other surrounding devices. In the first way services can be accessed directly with a P2P communication among devices. However this functionality is not exposed to components outside the F2F enabler. These IoT Services constitute internal functionality required for the correct F2F interaction detection inference. In the second way, services are accessed through a centralized architecture, making use of the SocIoTal Context Broker. In this situation the IoT Service provides information about an occurrence of a F2F interaction and contextual data about the particular inference.

Figure 21 shows how the information provided by the F2F enabler and exposed as IoT Service running on mobile phone is generated and integrated with other SocIoTal components.

When two SocIoTal devices, namely A and B, come into proximity, the type of their F2F relation is measured by making use of their Bluetooth (BT) radio and other mobile sensors to estimate distance and facing direction of the devices. The extracted raw information is fed to the Context Inference module and knowledge about social relation is extracted according to a defined SocIoTal Context Model (Figure 20). Such information can be shared with direct communication among the involved devices, for real-time and decentralized operations, but it is also shared with other central SocIoTal

components for remote access and use. By making use of the provided NSGI-9 (e.g., registerContext) interface, sharing of the extracted context is pushed by the Context Communication module on the device through the centralized SocIoTal Context Broker.

The shared social relationship estimation is accessed by the central *Trust & Reputation Manager* to infer about the trust relationship among people and compute reputation score. The trustworthiness of a device/user with respect to another can be initially considered as a weighted average (between 0 and 1) of the number and type of observed relations between two well-defined devices. As an example, personal relations can be weighted as 0.5, while social can weight as 0.3 and public as 0.2. Two given device/user couples tend to trust more each other if the number of personal relations they experience is higher. This allows to build a social graph derived from real social relations that can be used alone or in conjunction with existing ones (e.g., Facebook, Twitter, LinkedIn) to assess the degree of relations between two different devices/users. On the other hand, by combining all the social relation incurred by a device, a reputation score can be also built. The reputation score can be a weighted average (between 0 and 1) of all the number and type of relations observed by a given device, including also relations with (non)SocIoTal devices. As an example, personal relations can be weighted as 0.3, while social can weight as 0.2, public as 0.1 and those with (non)SocIoTal devices can be averaged as 0.4. By doing so, devices that tend to be in contact with many (non)SocIoTal devices receive a lower reputation scores as they can be in situation exposed to more security threats. When a change in the reputation score is detected such information is pushed back to the Context Broker using the NSGI-10 updateContext interface and accessed from interested components that can subscribe to the Broker by using the provided NSGI-10 subscribeContext interface.

More adequate ways to combine such information in order to account trustworthiness of relations and reputation score is investigated as part of WP2 work in Task T2.3.

Similarly the same information can be locally accessed by the *On Device Trust Manager* to internally compute reputation scores thus avoiding sharing any information with external components and preserving privacy. In addition the On Device Trust Manager can request reputation score of a given device (e.g., device B) requiring access to device provided services (e.g., data sharing) through the Context Broker and the central Trust & Reputation Manager and use it to verify the suitability for the device to access the required functionalities.

Additionally, the F2F enabler information generated by the device A is communicated to the *Authorization component* through the SocIoTal Context Broker. The provided information is used to define the context surrounding the device, in terms of devices and observed relations. Such information is used by the Authorization component to apply various rules about the services a device B (which eventually was previously in contact with device A) can request and access onto the device A. Such information can be used alone or together with the one provided by the Trust & Reputation Manager in order to issue a capability token. Such token is provided to device B using the Context Broker or dedicated direct interfaces and presented to device A, which verifies it using its *On Device Authorization Manager* before granting access or not to the required services.

While this approach assumes that data are consumed by accessing directly to the intended devices which provides the data, in case the shared information is accessed through the Context Broker, the F2F context information is used as attribute to annotate any sensing data provided by a given device and access to the Context Broker from requesting device B should be regulated by the centralized Authorization Manager in the same way.

8.1.4 Interfaces with privacy-preserving and context-sensitive communication blocks

While interaction and communication among components generating and consuming information provided by the F2F IoT service, can be locally achieved in the same way through the Context Communication module, in order to allow interaction of the F2F IoT service with local components deployed in SocIoTal target devices, namely smartphones, or in case of direct interaction with infrastructure blocks, a set of APIs has been defined and implemented.
As per D1.3.1 [3], the F2F enabler is accessed through a public API that includes the following four functionalities:

1. createF2FList (DeviceAddr). This function initializes the list for a particular device and returns all the F2F discovered devices that belong to the timestamp of the initial request.

2. readF2FList (DeviceAddr): [Device1, Device2, ..]. This function returns all devices that have F2F interaction with the device given as parameter that belong to the timestamp of the initial request.

3. updateF2FList (DeviceAddr). This function updates the list of the given device with the F2F discovered devices that belong to the timestamp of the initial request.

4. deleteF2FList (DeviceAddr). This function deletes the list of the given devices on the SOCIOTAL servers.

The required attributes to share context and information are presented below:

1. Local device ID. This is the ID identity of the device that performed the inference and detected the social interaction. For SocIoTal registered devices this ID is replaced by the pseudonyms provided by the SocIoTal identity manager.

2. Remote device ID. This is the ID identity of the device that was discovered and relevant information was exchanged in order to infer about on-going social interaction. For SocIoTal registered devices this ID is replaced by the pseudonyms provided by the SocIoTal identity manager. For unknown devices this ID is replaced by a unique pseudonyms that identifies a special class of non-SocIoTal devices.

3. Inference result. This is the result from the inference that was performed i.e., if there was a social interaction or not and the type (e.g., personal, social and public). In case of unknown devices, not running the F2F enabler and associated IoT service, as the detected relation can only estimate interpersonal distance, a special type of relation can be defined and account only for unknown devices estimated being at least within the public zone. Such information is provided to classify the device context. E.g., a place with many unknown devices can be classified as public, while one with many trusted devices can be classified as private. Then this information may be used to trigger different device behaviour.

4. Timestamp. This is the time and date of the occurrence of the social interaction.

5. (Optional) Duration. This is the duration of the social interaction. It is worth noted that this parameter could be calculated by the timestamp.

These functionalities can be locally accessed by on device services, such as data sharing and discovery to decide the type of data generated (from embedded sensors) and shared and the type of

discovery procedure to perform and other services to expose, accordingly to the detected context. Detailed interaction with the F2F enabler and other device services, through the above defined APIs will be detailed in following WP3 deliverables ([6] and [7]). Detailed interaction with the Group Manager will be also discussed in upcoming deliverable, by explaining how data sharing within group can be related to social relation information as extracted by the F2F enabler, by creating group based on social relations.

8.1.5 Enabler as a tool

As described before the F2F enabler will implement an IoT service suitable for smartphones that will extract type of social relations incurred by the device with other enabled SocIoTal devices. Such enabler will be implemented as black box and provided as a framework that can be either accessed through SDK implementing the above described APIs in order to extend other applications and integrating other components. On a simpler way, the information generated will be initially exposed and make available for other infrastructure components and developers through the SocIoTal context broker, thus simplifying and standardizing integration.

8.2 Radiolocation enabler

8.2.1 Fulfilment of SocIoTal requirements and needs

In SocIoTal, the radiolocation enabler is seen as a device-centric functionality, in the sense that both input radiolocation measurements and location-based outcomes are locally issued and (hopefully) specific to each capable device. However, at a higher level, it can still comply with both decentralized and centralized operating modes (e.g., in terms of intermediary location estimation and/or subsequent exploitation of the produced pseudonym outcomes). It can be exploited by several SocIoTal security and communication modules, as detailed follows.

From radiolocation to device-specific identity

The location information derived from wireless localization is one key feature contributing to forge, reinforce and control the identity of independent devices. It indeed reflects not only the unique physical insertion of each device in its geographical environment (e.g., through absolute GPS or WiFi-based positioning) but also the direct interconnections with other collocated fellows (e.g. relative distances between mobile neighbours over feasible short-range peer-to-peer links). Location-based pseudonyms, which are generated out of raw radiolocation measurements, processed location results, or a combination of the latter, can thus complement the public IDs of the involved devices for authentication or authorization purposes. But they can be also exploited to generate local device-specific secrets (e.g., equivalently to PUF), thus enabling to generate further cryptographic material (e.g., session key, seed for PRNG…).

From contextual identity reinforcement to trust and reputation

Location-based pseudonyms make use of context information related to the spatial occupancy and wireless connectivity of the devices at a given time. To some extent, they can also take part into the acquired context itself. Thus, bridging location-based identity and trust, they are primarily intended in authentication protection overlay mechanisms so as to prevent impersonation attacks. Being known uniquely by the legitimate devices (either through secure sharing or guess) but supposed unknown (or at least hardly inferable) by an attacker, they indeed contribute to ease or assist following authentication and authorization procedures, as detailed in the examples below. On the other hand, as such pseudonyms usually aggregate more complex information than just an explicit 2D position, they are expected to make the attacks (e.g., based on brute-force guess) much more challenging.

One step ahead, the “spatial identity” integrated over time can also be advantageously used to dynamically grant and refine “Trust & Reputation” (T&R) marks to mobile members or devices. For

instance, similarly to the Face-to-face enabler, this can be simply achieved by monitoring the space-time behaviour (and hence, the space-time consistency) between claimed, locally estimated/inferred and even learnt location-based information (e.g., raw radiolocation measurements, estimated locations or resulting location-based pseudonyms), trying to detect unexpected deviations.

8.2.2 Related IoT service

From the IoT service and architectural perspectives, as shown in the Figure below, the radiolocation enabler (i.e., location-based pseudonym generation) is intended for two modes:

- *Publish Data*:
  - (1) Location, as standalone information, is pushed towards LocationData Service and Location attribute of the VE is updated;
  - (2) The new locationData is notified to the pseudonym generation FC, which is a sub-component of Identity Management FC to generate a location-based pseudo after a subscription to the service;
- *Retrieve Data*:
  - A Security Functional Component like Authentication can request a location-based pseudonym directly to the Identity Manager (as shown on the figure below), or alternatively to the Context Manager (if location-based pseudonym are treated as part of the context information) with a request/response flow;

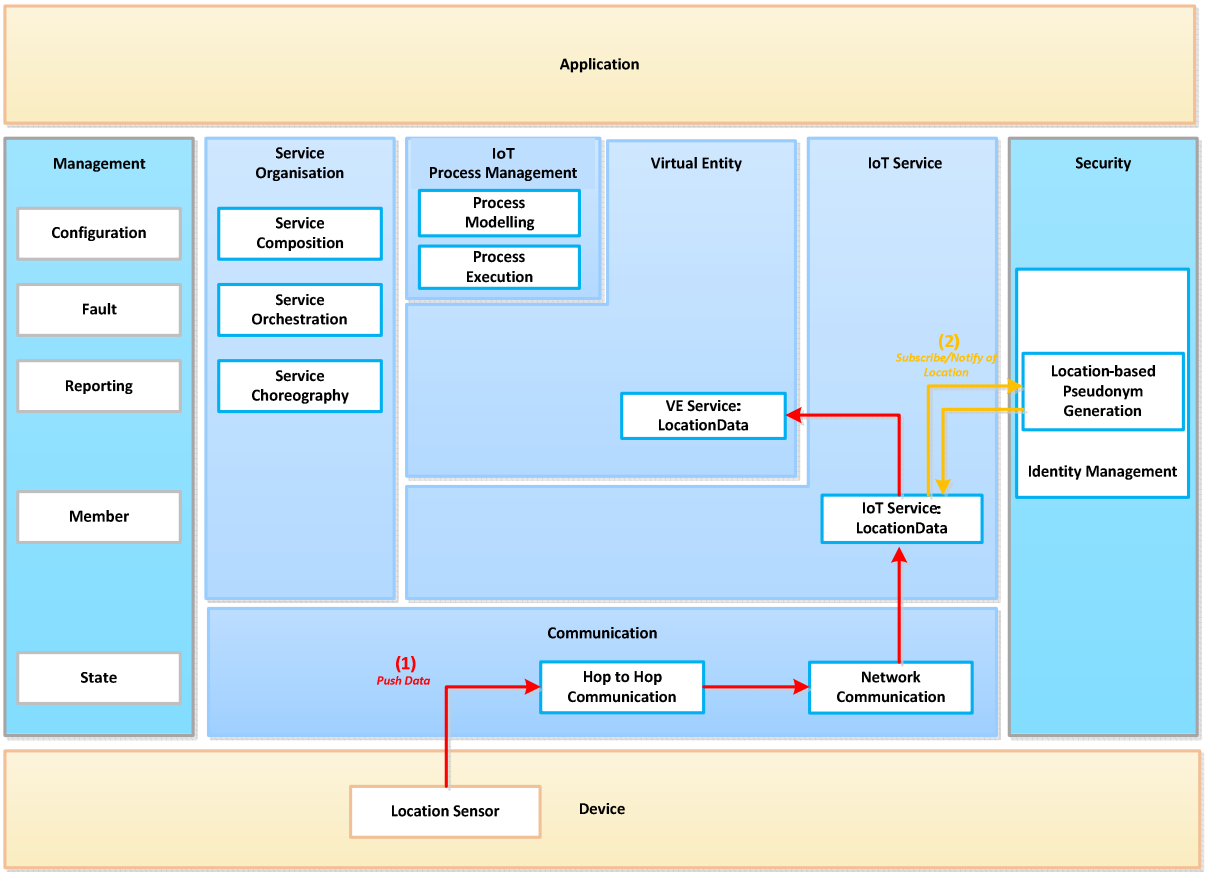

Figure 22. Radiolocation enabler (Location-based pseudonym generation) architecture compliant with IoT-A (functional/information view)

8.2.3 Interfaces with identity, trust and reputation management blocks

The figure below shows the most generic heterogeneous integration scenario involving the radiolocation enabler.

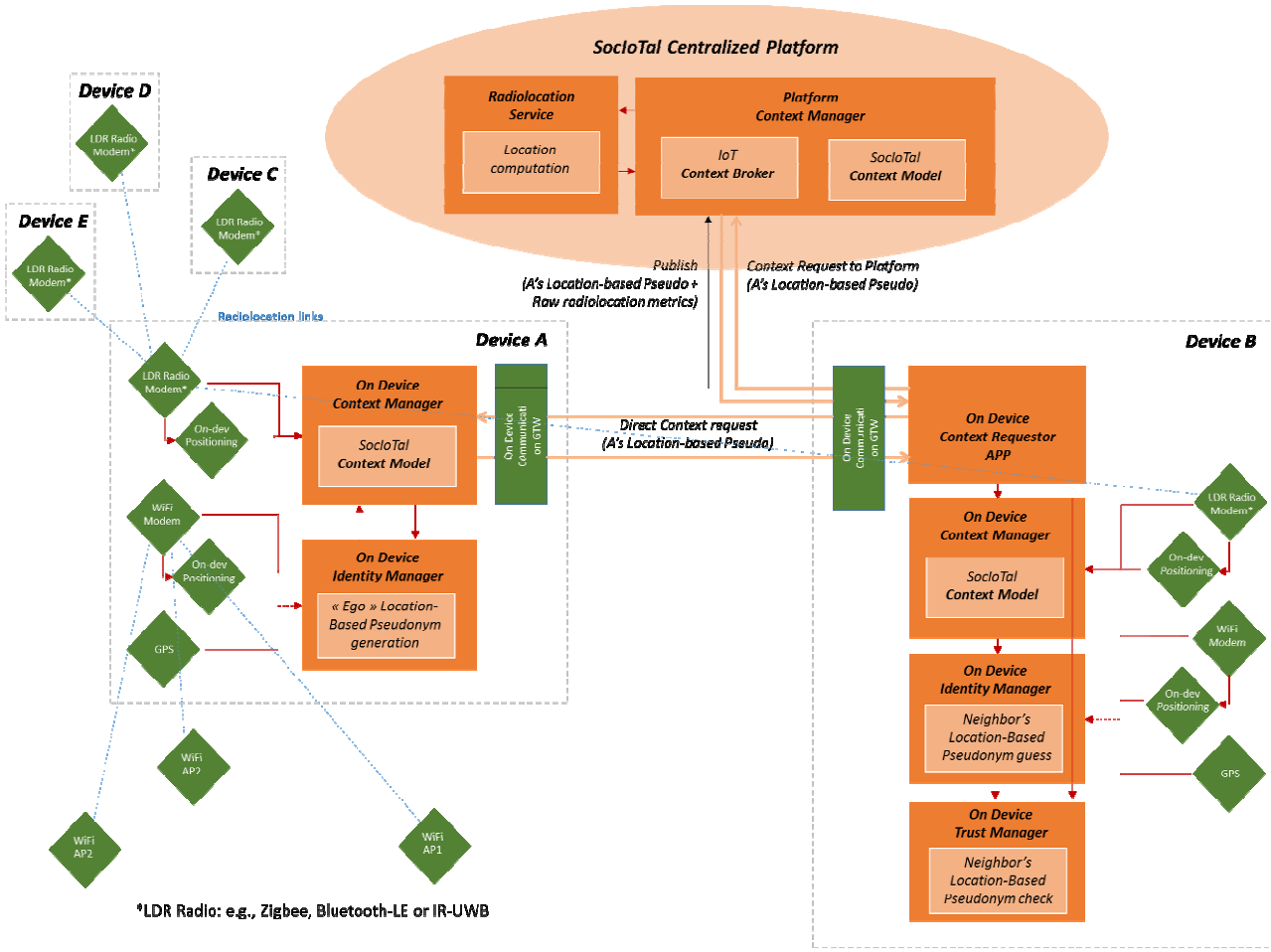

Figure 23. Generic integration scenario for the radiolocation enabler with context, identity and trust managers

Two integration examples in line with the high-level use cases illustrated in Sect. 3.2 are detailed below.

Ex. of context-based trust assessment through range-dependent pseudonyms verification (Impersonation prevention overlay)

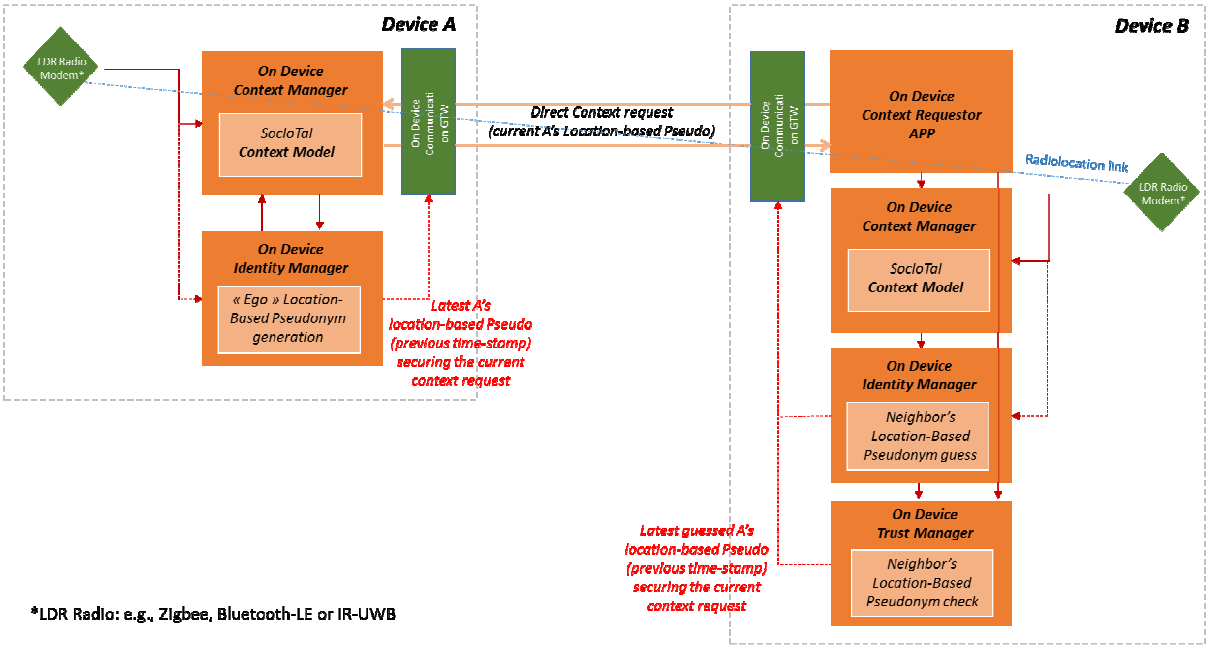

Figure 24. Particular integration example with context, identitiy and trust management modules enabling range-dependent pseudonym verification.

In this first example, given a couple of Devices A & B, Device A will first perform radiolocation measurements with respect to B (e.g., exchanging specific data packets through Low Data Rate radio means), which feed the on-device context manager (according to the SocIoTal context model), and even optionally, the on-device identity manager directly. The latter, which is in charge of elaborating locally Device A's location-based pseudonym (LBP), could also retrieve the required input information from the on-device context manager alternatively. Finally, it feeds back the on-device context manager with the computed LBP result (i.e., the LBP, which depends on both time and space, being viewed as part of the acquired context as well). Meanwhile, Device B acquires its own radiolocation measurements, feeding its own on-device context manager and one step further, its own on-device Identity Manager. The latter produces not only Device B's LBP, but it also makes one guess on Device A's LBP, based on its acquired measurements. This guess then feeds the On-Device Trust Manager. In parallel, Device B will directly and "securely" require the current LBP generated by A, through the on-device context requestor and the on-device communication module (e.g., using the latest available pseudonym associated with A on both sides of the link, as previously generated and guessed at the previous time stamp respectively by the two parties). The current LBP information claimed by A is then transferred to Device B's on-device trust manager for comparison purposes with the local guess generated by B. The time-stamp associated with the claimed pseudonym may be checked against a given validity timeout (e.g., seen as no more reliable if issued a too long time ago). Similarly to other enablers, in a steady-state regime (i.e., if a first secure link has been established at least once between B and A, after conventional authentication methods), this LBP cross-check procedure can then assist and complete any Trust assessment procedure based e.g., on location/proximity information, or a token-based authorization procedure.

Ex. of context-based trust assessment with position-dependent pseudonyms (Impersonation prevention overlay)

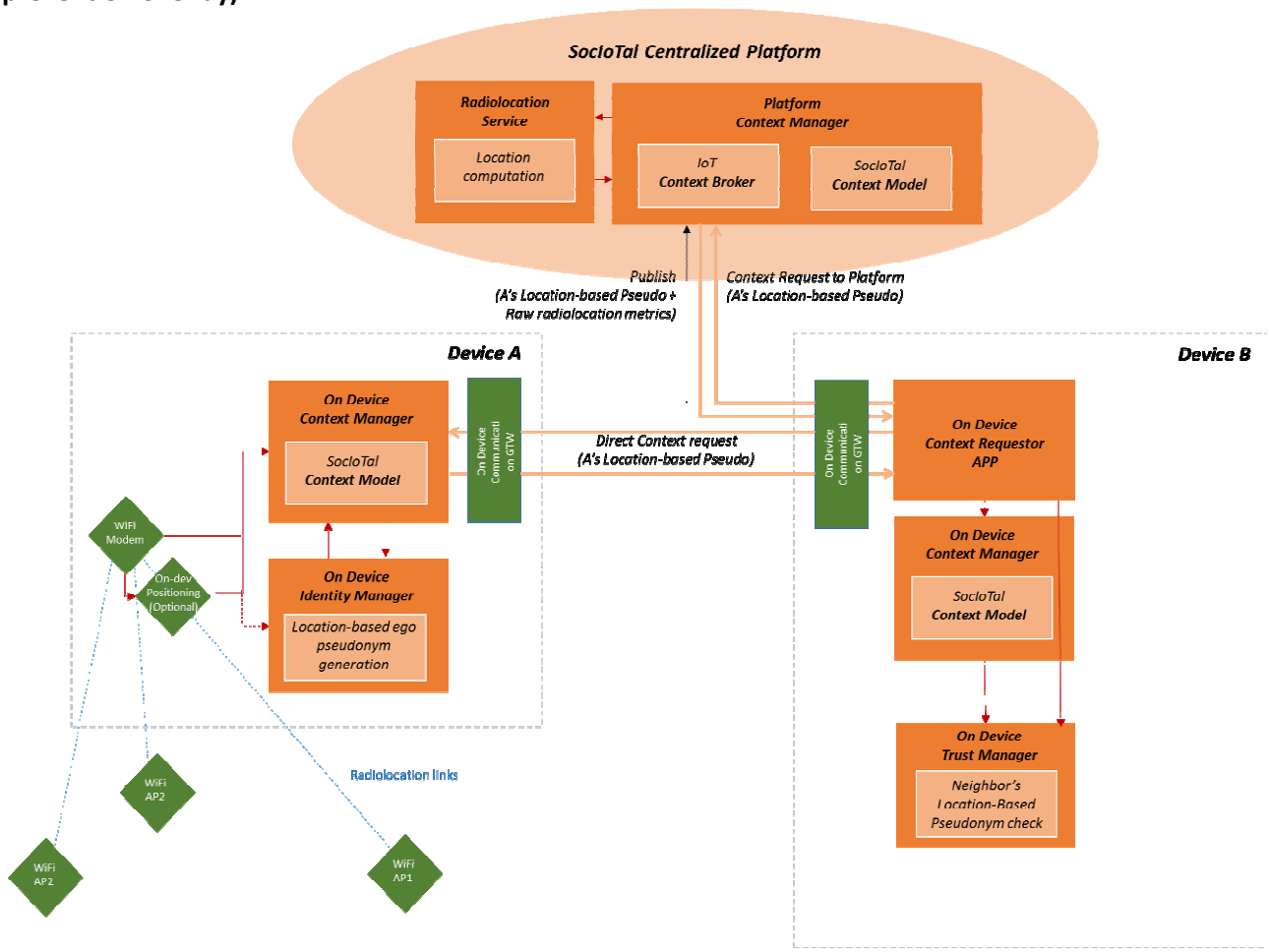

Figure 25. Particular integration example with context, identitiy and trust management modules enabling position-dependent pseudonym verification.

In this other example, the goal of Device B is still to compare and verify the LBP claimed by A but now with the corresponding LBP stored as contextual information about A by the centralized context IoT broker. Device A will first sense its radiolocation measurements and locally deliver (or get from a localization service) its positioning estimate, which feeds the on-device context manager (according to the SocIoTal context model), or even the on-device identity manager directly. The latter, which is in charge of elaborating locally Device A's location-based pseudonym (LBP), can also retrieve the required input information from the on-device context manager alternatively. Finally, the on-device identity manager feeds the produced LBP result back to the on-device context manager (i.e., the LBP, which depends on both time and space, being viewed as part of the acquired context). All this information is then published through the on-device context manager and communication module, and finally stored into a centralized database managed by the IoT context broker. The time-stamp associated with the claimed pseudonym may be also delivered (for further validation against a given validity timeout at B). Device B will then require Device A's LBP both i) from A directly and ii) from the centralized IoT context broker, through its own on-device context requestor and the on-device communication module. Both received elements are then transferred to Device B's on-device trust manager for comparison purposes. In this case again, in a steady-state regime (i.e. if a first secure link has been established at least once between B and A, after conventional authentication methods), this LBP cross-check procedure can assist Trust assessment or token-based authorization

procedures. So as to prevent next eavesdropping attacks (exploiting past LBP disclosure), right after delivering its LBP on demand to B, A should triggers again its on-device context and identity managers to restart the whole process and renew its local LBP based on refreshed connectivity information.

8.2.4 Interfaces with privacy-preserving and context-sensitive communication blocks

From a communication-oriented perspective, as already seen in sub-sections 3.2 and 8.2.3, beyond just providing a means to assist authentication procedures, the LBP enabler could be also beneficial to contextually protect communications between trusted users/devices sharing the same geographical area. For instance, LBPs can be used (and likely frequently renewed) to address the devices in direct data transmissions (instead of using public IDs or assigned MAC addresses), or they can even be used as (temporary) context-dependent common sources of secret, such as private keys in symmetric cryptography or session keys.
However, in both cases (i.e., pseudonym usage while addressing/sourcing data packets or basic input cryptographic material), links with respect to other communication blocks are clearly not enabler-specific (i.e., other kinds of pseudonyms could be intended, as well as other cryptographic material, while feeding exactly the same communication blocks). As such, these links will not be detailed in this subsection. Contrarily to authentication through LBP verification (which pre-supposes the existence of specific protocol transactions), the generated LBP data can at least be part of the context information exploited as input by such communication blocks (i.e., regardless of the way this content has been generated).

8.2.5 Enabler as a tool

Like F2F, the radiolocation-based enabler could implement an IoT service suitable for smartphones. But the generated information can be also made available through the SocIoTal context broker for simplicity. The required attributes to share related context information are presented below.

First of all, explicit context information related to location acquisition, in the most generic (i.e., heterogeneous) radiolocation scenario, could comprise some of the following data:

1. **{X_ i,Y_ i}** → Device i's estimated absolute 2D coordinates
    - directly issued at embedded GPS sensor or
    - computed out of WiFi RSSI-based fingerprinting w.r.t. surrounding Access Points (APs) or
    - computed out of Zigbee RSSI, Bluetooth-LE RSSI or IR-UWB RT-ToF measurements w.r.t. fixed Wireless Sensor Nodes (WSN) anchors/beacons and/or mobile neighboring devices, respectively through trilateration or cooperative estimation algorithms.
2. **T_COORD_i** → Time-stamp associated with the delivery of Device i's estimated absolute 2D coordinates;
3. **{RANG_ij, …}** → Set of single-link range measurements of node i w.r.t. neighboring nodes j (fixed WSN anchors/beacons and/or mobile neighboring devices).
    - Zigbee RSSI-based range measurements or
    - Bluetooth-LE RSSI-based range measurements or

- Impulse Radio - Ultra Wideband (IR-UWB) Round Trip – Time of Flight (RT-ToF) based range measurements.

4. **{RCD_ ij, …}** → (Optional) Set of relative relative clock drift measurements of node i w.r.t. neighboring nodes j (fixed WSN anchors/beacons and/or mobile neighboring devices), available only in case of side RT-ToF measurements acquisition.

5. **{RSSI_ij, …}** → Set of single-link WiFi RSSI measurements of Device i w.r.t. surrounding Access Points (APs) j.

6. **{T_MEAS_ij, …}** → Set of time-stamps associated with single-link measurements (IR-UWB RT-ToF –based ranges, Zigbee RSSI-based ranges, Bluetooth-LE RSSI-based ranges, or WiFi RSSI) of node i w.r.t. nodes j, the latter being fixed WSN anchors (beacons) and/or APs and/or mobile neighboring devices.

Then location-based pseudonym (LBP) information relying on the previous radiolocation sources, as part of the acquired context, can comprise the following data:

7. **DSLBP_i** → Device-specific location-based pseudonym generated for device i

8. **T_DSLBP_i** → Time-stamp associated with the device-specific location-based pseudonym generated for device i.

9. **{LSLBP_ij,…}** → Set of link-specific location-based pseudonyms (LBP) generated for device i w.r.t. neighboring nodes j.

10. **{T_LSLBP_ij,…}** → Set of time-stamps associated with the computation of link-specific location-based pseudonyms in use for device i w.r.t. neighboring nodes j.

The two exhaustive lists above mean neither that the devices will necessarily host all the previous radiolocation capabilities (i.e., GPS, WiFi, LDR radios) nor that these means will have to be operating simultaneously at a given device (e.g., one may want to save power consumption or limit interferences). Depending on the security strategy (i.e., given some risks assessment policy in a particular environment) and/or intended usage of the outcome information (e.g., for identity management, trust & reputation management, additional "protection" overlay assisting conventional authentication or authorization procedures…), only a subset might be available and combined.

8.3 Magnetic field localization enabler

8.3.1 Fulfilment of SocIoTal requirements and needs

From location to access control

As use case of the integration of location data for security mechanism, in the following we show an example of its applicability for distributed access control to smart objects.

This distributed access control is built on top of distributed Capabilities Based Access Control (CapBAC) [62], which is described in detail in the deliverable D2.2 of this project. This system makes use of an IAP-based communications architecture with emerging protocols which have been designed for constrained environments, such as 6LoWPAN or Constrained Application Protocol (CoAP).

An overview of our proposal to location-aware access control system is shown in Figure 26.

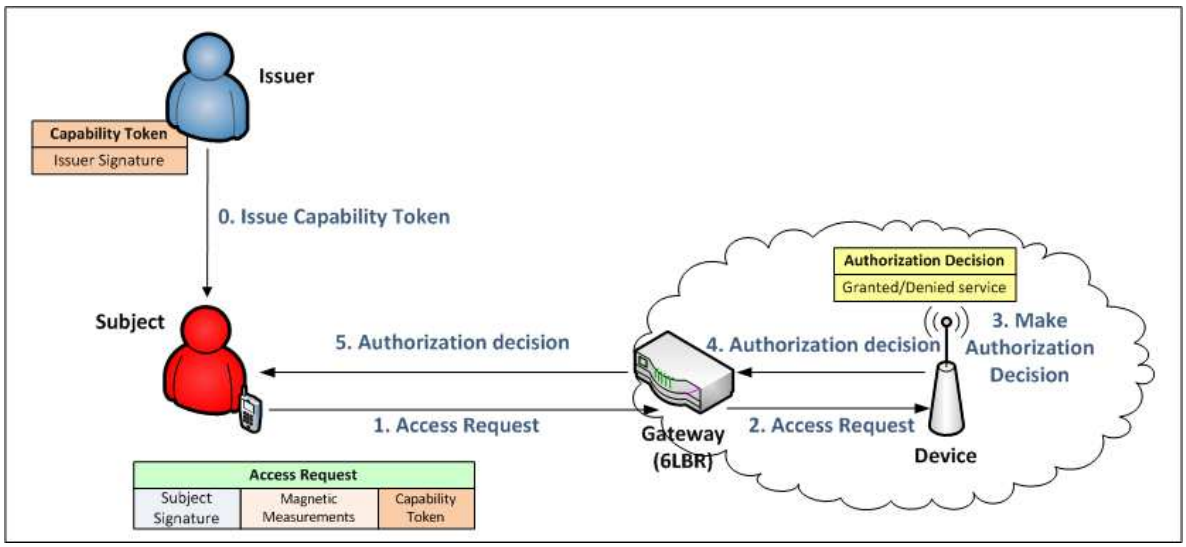

Figure 26. Proposed scenario for location-aware access control in buildings

The basic operation of our access control mechanism is as follows. As initial step, the issuer entity of the system, which could be the device's owner or manager, issues a capability token to the subject granting permissions on the device. Furthermore, such issuer signs this token in order to prevent security breaches. Once the subject has received the capability token, she tries to make use of the smart object. To do this, when she is close to the geographical area of the target device, she generates a request including magnetic field values and the capability token. In addition, this request must be signed in order to get access to the smart object. For this purpose, the CoAP request format has been extended with three headers: the capability, the signature and the magnetic field vector.

The first task to be performed by the authorization engine is the assessment of if the subject is inside the same building's zone where the smart object is placed. For this, we base on the magnetic field characterization associated to the landmark identified in such zone. Such landmarks' centroid is represented through a mean and deviation values associated to each magnetic field feature. Deviation is the parameter which indicates the zone's extension covered by each landmark in terms of magnetic field.

Therefore, given a device located in a building's zone where the magnetic field landmark lj with centroid Cj has been identified, the required device must assess if the distance between the mean values of the landmark centroid and the vector of magnetic field features extracted from the measurements sent by user, is smaller than the deviation associated to such landmark's centroid. If it is smaller, then it means subject is inside the same building's zone as device. Otherwise, the authorization process is aborted and the service is denied. If the previous requirement is satisfied, the second evaluation task is carried out. It consists of evaluating the capability token, which is attached to the access request. In case the capability token is successfully evaluated, the last task involved in our authorization engine is launched. During this step, it is evaluated if subject is inside the security zone defined for the required service (which can be denoted as SZ). For this evaluation, it is necessary to estimate firstly the subject position by using the RBF defined for the associated landmark. Once subject location is estimated Zk, the distance between subject and device is calculated, and then, it is evaluated if such distance is smaller than SZ. For this last evaluation, it is

considered the mean accuracy value (μz) associated to the RBF utilized to estimate the subject position.

A similar approach can be considered to other security mechanisms, such as for trust and reputation computation. Furthermore, this indoor localization mechanism is able to provide different levels of accuracy in its data depending on the granularity of the clustering applied during its off-line phase, in this way, different levels of computational cost and time consumption can be considered according to the final requirements of the security mechanism implemented.

8.3.2 Related IoT service: Indoor Localization

From the architectural point of view, Figure 27 shows the functional view for the indoor location enabler. The description of this view is as follows:

The Magnetic Device Sensor push data to the MagneticMeasurement IoT service. Here security consideration should be taken: the device makes a push operation over an IoT Service. So, an «IoT Service Push data security» flow should be performed according to the description in D1.2.1 [2].

The data is pushed in Indoor Location Detection VE Service, which calculates the indoor location of device A. (Notice that this VE Service may be placed in a device B different from where the MagnecticMeasurement IoT Service is placed)[2]. In case two devices were involved, here security consideration should be taken: the IoT Service tries to access a VE Service like a user or application tries to access a service. So, a «VE Service Request security» flow should be performed according to the description in D1.2.1 [2].

The quantified location is sent back to device A. and the VE Service updates the location position in the IoT LocationPosition Service. In case two devices were involved, here security consideration should be taken: the VE Service tries to access an IoT Service like a user or application tries to access a service. So, an «IoT Service Request security» flow should be performed according to the description in D1.2.1 [2]. Afterwards, the Indoor localization service can be used in different scenarios to feed the Context Manager Component of the framework and then make security decisions according to the localization context.

[2] In case two devices were involved, VE Resolution process as well as the flow through the Communication Functional group to communicate device A and B should be need. These flows are omitted in the diagram for the sake of clarity.

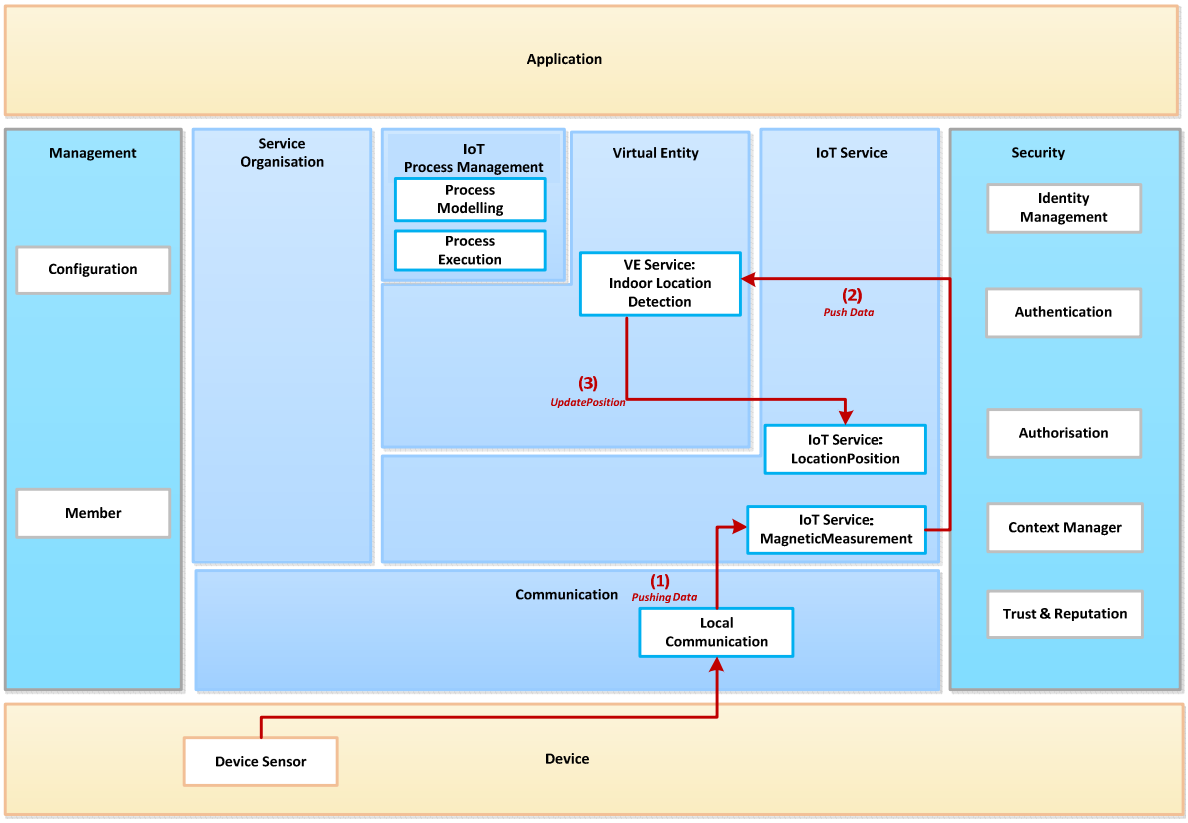

Figure 27. Indoor location enabler architecture compliant with IoT-A (functional/information view)

8.3.3 Interfaces with identity, trust and reputation management blocks

The main interaction of the localization enabler is within the Context Manager component of the SocIoTal security framework. The Indoor location enabler feeds the Context Manager with the obtained localization information. From the security perspective, the localization data obtained from the enabler is then delivered to the Context Manager. _Then the context is used mainly by the Authorization component of the framework to make authorization decisions accordingly.

The role of the Context Manager within the SocIoTal security framework is to provide the context notion to the different architecture modules in order to support their activities. The device Context Manager supports NGSI standard compliant and deal with the local context and be able to obtain context from the global Context Broker. The localization service can interact with the Context Manager by means of the NGSI interface.

The context can be managed globally in a backend and locally in devices or gateways. SocIoTal security components are able to made security decisions taking into account their local context along with the context coming from the backend. The Context Manager can publish raw context events (e.g., coming from the sensor) or elaborated events (as a result of the events processed by the context manager engine) to the backend context broker. The context broker is usually placed in the Cloud or in a Data centre in order to maintain and process context events coming from different devices. In addition to publish events, the Context manager can also accept events from any NGSI compliant Event Producer. Thus, the devices could be notified by the context broker with new context events needed to take security decisions.

Different security components, namely, the Identity Management, Authorization component, Group Manager and the Trust & reputation can be subscribed to the context engine and take security decisions depending on the inference of the rules defined in the Context manager inference engine. Thus, each time a rule is activated the security components can be notified with the consequent information derived in the rule. Figure 28 shows the way different security components of the security framework can access to the Context Manager to make security decisions accordingly.

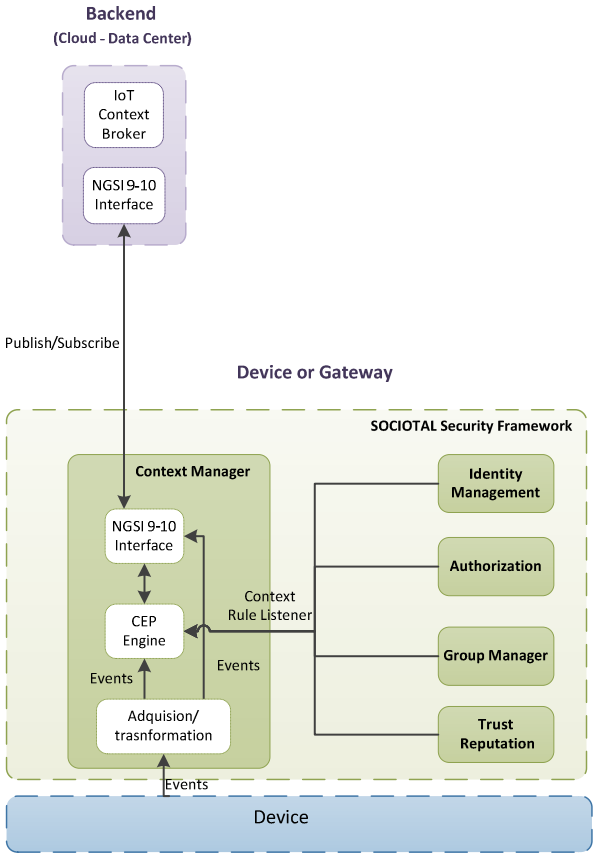

Figure 28. Context Manager main interactions with security components

The device Context Manager can be registered as an NGSI context provider in the backend IoT Context Broker, it can be done by calling the NGSI-9 registerContext operation. After context registration, the Context Manager can send events by calling the NGSI-10 updateContext method of NGSI interface in the backend. The Context broker acts as event consumer.

Additionally, the device Context Manager can be subscribed to receive context from the IoT Context Broker. It should be noted that the communication with the Context Broker could be subject to security mechanisms to ensure only trustworthy devices interact with it. Thus, the device can take security decisions locally based on the context when it is necessary. In this case, the Context Broker acts as context provider. The main security components, can access to the inferred context data in the local Context engine.

The Authorization functional component of the SocIoTal security framework is based on a combination of access control models and techniques. In order to accomplish with the main features of the proposed system, contextual information is a key aspect to be considered when making access control decisions. According to the SocIoTal access control system, a subject entity gets a capability

token in order to get access data from a target device. This token is usually generated by an Issuer entity which makes access control decisions that are embedded into the token. Therefore, when the subject entity tries to get access to a resource being hosted in a target entity, it provides the capability token previously obtained. Such token can contain contextual restrictions to be locally verified when the token is evaluated by the target device. At that moment, the target device can use contextual information from its local Context Manager, as well other data stemming from the IoT Context Broker deployed in the Backend. This process is shown in Figure 29, in which context information from the Context Manager is used by the target device when verifying the capability token.

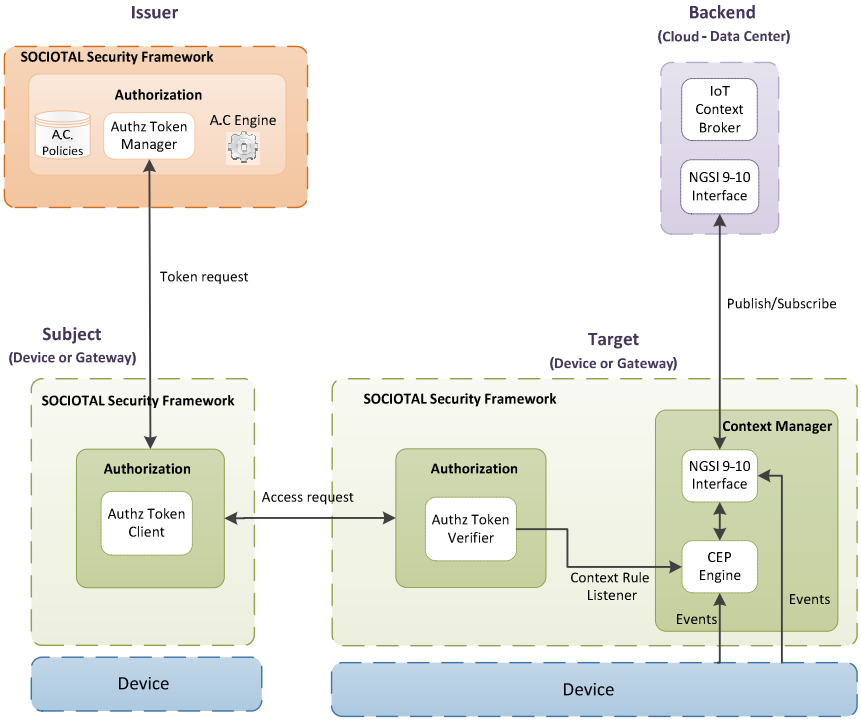

Figure 29. SocIoTal context-aware access control

The following snipped of code shows the indoor localization context Information provided by a SmartObject to the Central Context Manager regarding an interaction within another SmartObject. This context information is used to take the access control decision according to the location of the devices.

```
{
“contextElements”:[
    {
    type”: “urn:x-org:sociotal:resource:device”,
    “id”: “SocIoTal:hvacsystem:floor1.SmartObjectC.sensor.temperature,
    “isPattern”: “false”,
    “attributes”: [
        {
        “name”: “SensorIdentifier”,
        “type”: “http://sensorml.com/ont/swe/property/SensorIdentifier”
                    “value”: “SmartObjectX”
        },
        {
        “name”: “MagneticFieldMeasurements”,
        “type”: “MagneticField”,
        “value”: “{[0],[0],[0]}”
        },
        {
         “name”: “Location”,        //Localization coordinates
        “type”: “http://sensorml.com/ont/swe/property/Location””,
        “value”: “{0, 0}”
        },
    ]
}],
}
```

8.3.4 Interfaces with privacy-preserving and context-sensitive communication blocks

The outcome of the indoor localization enabler can be used by the Group Manager component in order for it to take the localization context into account when sharing data securely. The sharing policies can be driven by the indoor localization context produced by the enabler and handled by the indoor Context Manager. The Group Manager component of the SocIoTal security framework is based on the use of the Ciphertext Policy Attribute Based Encryption (CP-ABE) [63] cryptographic scheme in order to enable a secure data sharing mechanism with groups of entities (i.e., communities and bubbles of smart objects. The following figure shows how contextual information from the Context Manager of a subject entity (acting as an information producer) is used by the Group Manager to select a specific CP-ABE policy to encrypt a specific information. Specifically, the Group Manager contains a set of Sharing Policies how the information is disseminated according to contextual data. These policies are evaluated by the Group Sharing engine before information is disseminated by the subject. The result of the evaluation of these policies is, in turn, a CP-ABE policy, which is employed by the CP-ABE engine to encrypt the information with that policy. After the information is encrypted and disseminated, the Group Manager of a target entity (acting as a consumer) will try to decrypt it with CP-ABE keys related to its identity attributes through its CP-ABE engine.

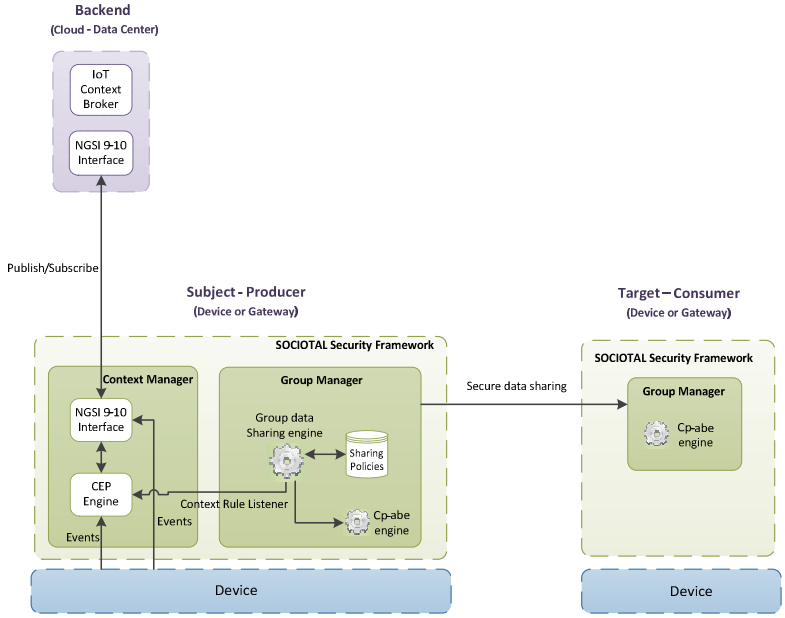

Figure 30. SocIoTal context-aware group sharing

Figure 30 shows an overview of the main components of the Security Framework required when the secure group sharing takes place between a producer and consumer. As can be seen, in the producer side the Group manager evaluates the sharing policies taking into account the Context obtained from the indoor localization enablers and handled by the local Context Manager, which in turn, can interact with the Context broker. As a result of the sharing policies evaluation, a CP-ABE policy is selected to encrypt the data, which is then delivered securely to the consumer(s).

8.3.5 Enabler as a tool

The magnetic field localization enabler will implement the indoor localization service. This service, as it has been described before, is able to determine the location of a subject device analysing the magnetic field measurements. The service can be accessible by any device connected to the Internet. Additionally, the localization measurements will be exposed and make them available to the Context Broker, which will make the localization even more accessible for other infrastructure components.
Additionally, in order to simplify the integration process and deal with the indoor localization access control scenario, an android app will be implemented to enable smartphones to access the indoor localization service to obtain the location of the device being analysed.

Section 9 - Enablers security

The methodology for risk analysis that has been chosen to assess the security/privacy of device-centric enablers is based on Microsoft STRIDE / DREAD, like in D1.1 [1].

The following elements to be protected in case of device-centric enablers were identified:

- **Physical person**
- **Subject's privacy** (user or device)
- **Communications channel**. (e.g., alteration of exchanged data integrity through tampering and replay attacks, alteration of routing functionality through black hole, worm hole, depletion...).
- **Leaf devices** (e.g., integrity of the software, hardware, and location).
- **Intermediary devices** (e.g., Disabling or tampering a gateway for denial–of-service).
- **Backend services** (e.g., Compromising server-side applicative software and data).
- **Infrastructure Services** (e.g., Altering Discovery, Lookup and Resolution Services, or Security Services such as Authorization, Authentication, Identity Management, Key Management and Trust and Reputation).
- **Global service**

According to STRIDE classification, the risk sources can be as follows:

- **Identity spoofing** (one peer illegitimately uses the identity of another peer).
- **Data tampering** (one attacker alters the content of data exchanged between legitimate peers).
- **Repudiation** (one attacker performs illegitimate action but denies having performing it in such a way that no one can tell).
- **Information disclosure** (information is disclosed to unauthorized peers).
- **Denial of service** (a service offered to legitimate users is denied).
- **Elevation of privilege** (in systems featuring different classes of users with specific rights, one attacker manages to acquire rights that would normally be granted to more privileged class(es).

We define in the table below intersections between STRIDE items and elements to protect into generic identified risks.

		Risk family/source					
		Spoofing Identity	**Tampering with Data**	**Repudiation**	**Information Disclosure**	**Denial of Service**	**Elevation of Privilege**
Elements to protect	**Physical person**		Attack alters data so that wrong data is supplied to a critical monitoring system	Human Users might use unattended electronic devices leaving no digital trace		A service critical for user's safety is disabled	
	User's privacy	User's identity is spoofed User is involved in transactions with a malicious peer			Attacker gains knowledge of user private parameters Attacker gains knowledge of user's location		

	Communication channel		Alteration of the invocation of a service Alteration of the return value upon service invocation	Jamming wireless communication channels lead to local DoS attacks that can be repudiated	Attacker gains knowledge of sensitive exchanged data	Attacker disrupts communications	Wrong authorization information propagating from one server to another
	Leaf devices	Loss or theft of physical device used for authentication Attacker changes the association between a Virtual Entity and the corresponding Physical Entity	Attacker gains control of an actuator Attacker alter leaf device content so that a user will eventually be redirected to a malicious content Attacker alter sensor device so that monitoring of a Physical Entity fails		Disclosure of device configuration information Device identification Loss or theft of physical device containing private information	Attacker physically disables leaf device (local) Attacker physically disables leaf device (remote) Attacker prevents proper communication to an actuator	
	Intermediary devices		Compromised intermediary devices alter traversing data	Intermediary devices behave maliciously and clients are not able to report the fact		Assisting intermediary devices are no longer usable	
	Backend Services	Administrator role usurpation Backend account hacked			Massive disclosure of collected data	Backend Service is made unavailable	
	Infrastructure Services	Attacker impersonates infrastructure Services, compromising IoT functionalities and/or other dependent infrastructure Services	Attacker poisons infrastructure databases or alter outgoing information		Disclosure of private Services (existence & description) Disclosure of access policies Disclosure of Identities and cryptographic material	Attacker denies legitimate users access to Infrastructure Services	
	Global systems / facilities				Massive disclosure of users personal information	Disruption of a global service	

Table 6. Generic risks identification

Identified risks can be assessed for each enabler using the DREAD methodology & metrics, rating **D**amage Potential, **R**eproducibility, **E**xploitability, **A**ffected Users, and **D**iscoverability on a simplified scale {L (Low), M (Medium), H (High)}. The focus is mainly on the direct consequences to the enabler it-self. Further analysis would be required for services and management blocks based on the enablers under attack.

9.1 Face-to-face enabler security

The risks jeopardizing or threatening the application of face-to-face enablers are listed below, along with possible mitigation means (i.e., native or requiring slight modifications).

Risk family/source	Specific realization	D/R/E/A/D rating	Immediate implication to enablers	Mitigation
Identity spoofing	User's ***public*** identity is spoofed in ***active peer-to-peer ranging*** transactions User is involved in ***relative ranging transactions*** with a malicious peer	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/M/M/M	An attacker overhearing the exchange of information between two legitimate enablers could collect information about device identity, i.e., Bluetooth id.	Legitimate SOCIOTAL devices, pre-emptively registered on the platform, could obtain pseudonyms and use them to prevent user and device identity. In addition communication could be encrypted.
Information disclosure	User is involved in ***relative ranging transactions*** with a malicious peer User is involved in ***relative ranging transactions*** with a legitimate peer and MiM attacker	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability M/L/L/M/M	Malicious user running the F2F enabler could obtain information about other SOCIOTAL device user's facing direction, interpersonal distance and coarse-grained location information. MiM attackers could obtain information about other SOCIOTAL user's facing direction, coarse grain location of surrounding devices.	Communication encryption should help to prevent MiM attacks, thus requiring to register all the SOCIOTAL enabled devices. Discovery and isolation of malicious user can be hard to be detected and prevent from the F2F enabler usage perspective, but it should be addressed in other ways, by identifying malicious behaviour following the gained information, such as anomaly detection.
Data tampering	An attacker could intercept packets between two legitimate peers and change their content. A malicious user could force and inject the estimation of different facing direction and interpersonal distance (i.e. change transmission power of Bluetooth	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/L/M/M	The different peers affected by the attack might be puzzled and be unable to estimate proper F2F interaction or estimate wrong ones.	As the peers are expected to receive both tampered and genuine packets and information it should be easy to naturally detect anomalous situation and filter them. History of previously detected F2F interaction between involved peers could help to identify potential data tampering situation. In addition as the distance estimation is performed at device level, wrong facing direction could be filtered out as well, thus enabling to filter wrong F2F interaction detection.

	interface) information.			
Repudiation	A malicious peer could generate false direction information, thus generating wrong F2F interaction. Impersonation attack from a third party attacker could lead to infer wrong F2F interaction.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/L/M/M	By generating wrong information, a malicious user or an attacker could force wrong F2F estimation to be detected by peer device and their F2F enabler or be assigned to wrong device.	It could be difficult to detect such alterations and prevent from them simply looking at the F2F enabler behaviour. Detection will require analysing behaviour of peers. In addition trusted platform manager should be implemented in order to avoid tampering of direction estimation information. Frequent pseudonyms generation could mitigate impersonation attack.
Service denial	Being the F2F enabler completely distributed such risk shouldn't affect the system. Unavailability of Authentication server should be detected and handled outside of the enabler.			
Privilege elevation	A malicious peer could force the estimation of wrong F2F interaction	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/L/M/M	A malicious peer could generate wrong direction estimation information thus being able to force detection of wrong F2F interaction and compromising the system, by gaining access to privileges based on trust and device proximity.	Trusted platform manager on the device could prevent to compromise the generation of false direction and interpersonal distance information.

Table 7. Specific risks and mitigation for face-to-face enablers

9.2 Radiolocation enabler security

The risks jeopardizing or threatening the application of ad hoc radiolocation-based enablers are listed below, along with possible mitigation means (i.e., native or requiring slight modifications). Depending on the operating mode and device capabilities, part of the input information used in pseudonym generation can be i) external to the device it-self (i.e., extrinsic) or require external infrastructure resources (e.g., in case of centralized absolute 2D positioning localization requiring fixed anchors/Aps and/or centralized fingerprinting database like and/or centralized location computation server) or ii)

decentralized, staying local to the device it-self (e.g., knowledge of peer-to-peer connectivity and/or relative ranging measurements). This directly impacts the kind of possible attacks.

Risk family/source	Specific realization	D/R/E/A/D rating	Immediate implication to enablers	Mitigation
Identity spoofing	User's ***public*** identity is spoofed in ***active peer-to-peer ranging*** transactions User is involved in ***relative ranging transactions*** with a malicious peer	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/M/L/H	While acquiring ranging information over the illegitimate link, attackers may capture part of the input material used in legitimate nodes pseudonym generation, thus easing brute-force guess and further impersonation attacks based on illegitimately guessed pseudonyms (e.g. worst case within range-based pseudonym generation over single links). Erroneous or impractical pseudonyms generation could be forced at the device under attack (through injection of packets harmful to the ranging procedure) for further DoS of pseudonym-based services or transactions w.r.t. legitimate neighbours	In a steady-state regime, side monitoring mechanisms can be implemented to track and check the spatial consistency of input parameters (and mapping with public ID) w.r.t. legitimate neighbours (e.g., ranging history over time, checking compliance with plausible mobility assumptions).
Information disclosure	Attacker gains knowledge of ***final*** user's private ***location-based pseudonym*** parameter	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/H/L/L	Direct impersonation attacks are made possible against pseudonym-based data communication & security services.	Explicit disclosure over a public channel (even within a short duration of time) of the complete location-based pseudonym is not the preferred sharing option. Legitimate guess (even with moderate success rates) is highly favoured instead. Public leakage of non-sufficient building input material (involved in pseudonym generation) is also limited to the minimum. Location-based pseudonyms are used as protection overlay (but not necessarily unique means) into security schemes like authentication

Information disclosure	Attacker gains knowledge of user's ***absolute*** location ***through brute-force guess (in the operating zone), statistics-aided guess (e.g., prior knowledge of users' mobility habits) or hacked external positioning server***	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability M/L/M/M/L	For pseudonyms generated out of 2D positions only (extrinsic information), attacker would be able to generate pseudonyms similar to the legitimate device. Otherwise, brute-force guess can be eased at least.	Joint use of various sources of input information for pseudonym generation is the preferred option: device-specific information (relative clock drifts caused by HW clock imprecisions), neighbourhood-specific information (relative peer-to-peer distances) and possibly absolute location information.
Information disclosure	Attacker gains knowledge of user's public parameters (ID, MAC@…)	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability L/H/L/H/L	Attacker detains one of the most trivial ingredient feeding the pseudonym generation block, possibly easing brute-force guess again. Attacker may claim the legitimate public ID to launch ranging transactions with peers → see identity spoofing attack before	The public information is by far non-sufficient to forge the required pseudonyms, but just usable to actively trig/participate into ranging transactions.
Data tampering	Attacker gains control of a legitimate neighbour so that the content of the packets involved in peer-to-peer ranging transactions can be altered or manipulated	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/M/L/H	Erroneous pseudonyms can be forced at legitimate nodes due to false input information, causing possible DoS or service miss-operating.	In a steady-state regime, side monitoring mechanisms can be implemented to track and check the spatial consistency of input parameters w.r.t. legitimate neighbours over time.
Data tampering	Attacker compromises intermediary devices **or Gateways (e.g., Wifi AP)** to alter traversing data ***upon centralized localization service invocation*** Attacker poisons infrastructure databases **(e.g., RSSI fingerprinting**	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability M/M/L/M/H	For pseudonyms generated out of 2D positions only (extrinsic information), erroneous or impractical pseudonyms can be forced at legitimate nodes due to false input information, causing possible DoS or service miss-operating.	Generally, joint usage of various sources of input information for pseudonym generation is again the preferred option. Beyond, if a "centralized" 2D positioning source (e.g., centralized localization server) is not the only one available, it is preferable to favour local/decentralized input information instead (e.g., peer-to-peer ranges…).

	database) or alter outgoing information ***(e.g. result of a location computation at a centralized localization server) upon centralized localization service invocation*** Attacker alters the invocation of a ***centralized localization*** service or the return value upon ***centralized localization*** service invocation			
Service denial	Intermediary devices ***like Gateways/ anchors (e.g., Wifi AP)*** or back-end infrastructure ***(e.g. centralized fingerprinting database of centralized localization server)*** are no longer available upon ***centralized localization*** service invocation	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability M/M/L/H/H	For pseudonyms generated out of 2D positions only (extrinsic information), direct DoS for pseudonym-based services or data transaction.	Generally, joint usage of various sources of input information for pseudonym generation is again the preferred option. Beyond, if a "centralized" 2D positioning source providing extrinsic information (e.g., centralized localization server) is not the only one available, it is preferable to favour local/decentralized input information instead (e.g., peer-to-peer ranges…).
Service denial	Attacker disrupts ***peer-to-peer ranging transactions***	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/L/H/H	For pseudonyms generated out of peer-to-peer ranging, direct DoS for pseudonym-based services or data transactions.	One can temporarily switch back to simpler input information like 2D position in case of repeatedly failed ranging transactions (and thus suspected disruption).

Table 8. Specific risks and mitigation for radiolocation-based identity enablers

9.3 Magnetic localization enabler security

The risks jeopardizing or threatening the application of magnetic localization enablers are listed below, along with possible mitigation means (i.e., native or requiring slight modifications).

Risk family/source	Specific realization	D/R/E/A/D rating	Immediate implication to enablers	Mitigation
Identity spoofing	Subject and target users, or indoor location service identities are spoofed.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/L/M/M/H	Attackers could impersonate subjects obtaining services from target users. Additionally, attackers could impersonate target users providing fake services to subjects. Furthermore, the indoor location service could be impersonated providing forged location data.	Suitable cryptographic mechanisms for mutual authentication should be implemented for each transaction. The use of digital certificates guarantees identities cannot be spoofed.
Information disclosure	An attacker could obtain magnetic field measurements directly from the subject's smartphone, access rights contained in the capability token, or location data from the indoor location service	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/M/M/M/M	Location information could be faked by an attacker receiving magnetic field measurements. Furthermore, privacy of users can be threatened due to location information and access rights disclosure.	Confidentiality of communications involved are required through encryption. Location information could be protected in order to avoid disclosure to unauthorized third parties.
Data tampering	An attacker, acting as a subject user, tampers magnetic field measurements to convince he is in the security zone of the target user.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/H/H/M/M	The inference process of data location by the location service would be false in the case of magnetic field from the subject device are tampered. Target users could believe a subject user is in his security zone when he is not, and viceversa.	The indoor location service can ask a trust manager in order to get the trust score associated to the subject device. In turn, the Trust manager could require information from other devices or users (with trust scores associated), in order to infer if magnetic field measurements being provided by the subject device are have enough credibility.
Data tampering	An attacker, in the middle of subject and target users, tampers data being exchanged.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/H/H/M/M	Target users could believe a subject user is in his security zone when he is not, and viceversa. Furthermore, in the case of tampered data from the target user, service provisioning would be faked.	Suitable cryptographic mechanisms and protocols are required for secure communications.

Data tampering	An attacker in the middle of target and location service tampers location data being exchanged.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/H/H/M/M	Target users could believe a subject user is in his security zone when he is not, and vice versa.	Suitable cryptographic mechanisms and protocols are required for secure communications.
Repudiation	An attacker, acting as a subject user, tries to perform an illegitimate action over a target user, in such a way that no one can tell.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/M/M/M/H		Suitable cryptographic mechanisms and protocols are required for secure communications.
Repudiation	An attacker, acting as a target user, tries to perform an illegitimate action over the indoor location service, in such a way that no one can tell.	**D**amage Potential/ **R**eproducibility/ **E**xploitability/ **A**ffected Users/ **D**iscoverability H/M/M/M/H		Suitable cryptographic mechanisms and protocols are required for secure communications.
Service denial	A target user denies a specific action to an authorized subject user.	M/M/M/M/M	Service being provided by the target device is not given to the subject user.	The use of capability tokens is intended to avoid this risk.
Service denial	The indoor location service denies a location request to an authorized target user.	M/M/M/M/M	Location data are not provided to the target user.	The use of capability tokens is intended to avoid this risk.
Privilege elevation	An attacker, acting as a subject user, without a required privileged, is able to perform a specific action over the target device.	H/M/M/M/H		The use of capability tokens is intended to avoid this risk.
Privilege elevation	An attacker, acting as a target user, without a required privileged, is	H/M/M/M/H		The use of capability tokens is intended to avoid this risk.

	able to perform a specific action over the indoor location service.			

Table 9. Specific risks and mitigation for magnetic localization enablers

9.4 Overall assessment of enablers security

In the previous subsections, the security of each enabler has been assessed independently, by

- Identifying the preferred enabler-specific attacks (i.e., damaging or exploiting -in an opportunistic way- the enabler functionality);
- Evaluating the corresponding risks according to a standardized methodology;
- Suggesting tangible solutions to mitigate the previous risks, including:
    - Mechanisms that are intrinsic to the enablers themselves (e.g., considering heterogeneous input data -rather that a unique modality like a 2D position or a single-link relative distance-, while generating location-based pseudonyms, using intermediary F2F features like relative distance and heading to detect faulty interactions, monitoring the history of the input metrics feeding the enablers…)
    - Other existing/conventional techniques compatible with the enablers (e.g., cryptographic protocols, ciphering, bootstrap authentication…).

Overall, it should be clearly noted that:

- Security and privacy in the SocIoTal framework are treated following a holistic approach. Accordingly, the described device-centric enablers represent elementary building blocks, contributing to higher-level (end-to-end) mechanisms in charge of assessing trust and reputation on the one hand, and ensuring secure/private communications on the other hand (See e.g., integration examples in Section 8). For instance, trust management via location-/mobility-based anomaly detection and/or adjusting the levels and circles of trust based on users feedback could be beneficial to the enablers security (e.g., filtering out faulty F2F interactions) but dually, enablers would also contribute to feedback the trust and reputation manager with time-stamped data (e.g., dynamic locations, relative social graphs…).
- The proposed device-centric enablers are also naturally synergetic and mutually beneficial from a security point of view. Most often, they can be viewed as protection overlays to each other (e.g., position-based pseudonym validity can partly rely on the relative connectivity graphs discovered/maintained by F2F and location-based access/authorization can be granted based on the joint verification of physical locations, capability tokens and location-based pseudonyms…).

Section 10 - Conclusion

This deliverable represents the second and the last one released in reference to the work carried forward in Task T3.1 about device centric enablers. Therefore it concludes the work of the task. According to this, a number of additional contributions can be observed with respect to the previous deliverable.

All the enablers have been extended in terms of functionalities, and their performance evaluation has been improved by including new scenarios, evaluations of such novel functionalities and comparison against new benchmarks.

For instance the introduction of distance estimation, making use of Bluetooth RSSI measurements, has been presented and evaluated for the F2F enabler. This provides the enabler with a better and more accurate estimation of social relations.

The radiolocation enabler has been improved by using explicit P2P RSSI reading as input measurements so as to feed concrete range estimators before quantization, making the proposed solution less specific and more flexible. This enabler has been evaluated not only with simulations but also using real measurements.

Finally, an improved magnetic field based localization algorithm has been proposed thus improving the accuracy of the enabler in estimating indoor location in building with numerous magnetic interference sources. Results showed the better-achieved performance in real world scenarios, against selected state of the art benchmark based on WiFi localization estimation.

In addition a complete integration of all the enablers and way the information they provide will be exploited by other components of the SocIoTal platform has been discussed. This highlights the way the enablers will be integrated with WP2 components and their functionalities exposed to support development in other WP3 tasks, namely T3.2 and T3.3.

The evaluation and the provided integration specification show that the designed components are now mature enough to be integrated and tested in large trials, which will aim to evaluate also their implementation effectiveness in terms of energy efficiency and resilience to security and privacy attack.

For this reason, a first risk assessment of the different enablers has been performed showing how the integration of the different components into the SocIoTal platform will be beneficial in increasing their level of security and privacy. A more detailed evaluation in real scenarios where the enablers will be exposed to real security threats will be then performed in WP5, which will look at the holistic evaluation of the overall SocIoTal platform.

According to the level of proposed solutions and results achieved by Task T3.1 and assessed with the current deliverable, it can be concluded that the task overachieved its objective and provided solid innovative technologies from the integration of which, the other tasks in the WP3 and other work packages will surely benefit.

Section 11 - References

[1]. "SocIoTal scenarios and requirements definition report", Deliverable D1.1 of the SocIoTal project, Feb. 2014.

[2]. "First Version of SocIoTal Architecture", Deliverable D1.2.1 of the SocIoTal project, Aug. 2014.

[3]. "First version of API specification", Deliverable D1.3.1 of the SocIoTal project, July 2014.

[4]. "Framework Specification for Privacy and Access Control", Deliverable D2.2 of the SocIoTal project, Nov. 2014.

[5]. "Device centric enablers for privacy and trust (Intermediary)", Deliverable D3.1.1 of the SocIoTal project, Aug. 2014.

[6]. "Privacy-aware context-sensing device discovery", Deliverable D3.2.1 of the SocIoTal project, May 2015.

[7]. "Secure Group Communication", Deliverable D3.3 of the SocIoTal project, May 2015.

[8]. "Trials and pilots specification", Deliverable D5.1 of the SocIoTal project, Jan. 2015.

[9]. "SOCIOTAL evaluation", Deliverable D5.2 of the SocIoTal project, Sept. 2015.

[10]. J.J. Nielsen, et al. "Assessment of Cooperative and Heterogeneous Indoor Localization Algorithms with Real Radio Devices," in Proc. IEEE ICC'14, ANLN Workshop, Sydney, June 2014.

[11]. A. Conti, et al. "Network Experimentation for Cooperative Localization," IEEE JSAC, vol. 30, pp. 467–475, Feb. 2012.

[12]. A. Conti, et al. "Experimental characterization of diversity navigation," Systems Journal, IEEE, vol. 8, pp. 115–124, March 2014.

[13]. Z. Sahinoglu, S. Gezici, and I. Guvenc, "Ultra-Wideband Positioning Systems: Theoretical Limits, Ranging Algorithms, and Protocols," Cambridge Univ. Press, 2008.

[14]. D. Macagnano, et al. "MAC Performances for Localization and Tracking in Wireless Sensor Networks," in Proc. WPNC'07, pp.297-302, Hannover, March 2007.

[15]. M. Maman, et al. "Synergetic MAC and Higher Layers Functionalities for UWB LDR-LT Wireless Networks," in Proc. IEEE ICUWB'08. vol.3, pp.101-104, Hannover, Sept. 2008.

[16]. M. Laaraiedh, S. Avrillon, and B. Uguen, "Enhancing Positioning Accuracy through Direct Position Estimators Based on Hybrid RSS Data Fusion," in Proc. IEEE VTC-Spring'09, Barcelona, April 2009.

[17]. H. Cho, et al. "Performance Analysis of Location Estimation Algorithm in ZigBee Networks using Received Signal Strength," in Proc. IEEE AINAW'07, vol.2, pp. 302-306, May 2007.

[18]. K. Zeng, K. Govindan, and P. Mohapatra, "Non-Cryptographic Authentication and Identification in Wireless Networks", IEEE Wireless Communications, vol.17, no.5, pp.56-62,

Oct. 2010.

[19]. D.B. Faria and D.R. Cheriton, "Detecting Identity-Based Attacks in Wireless Networks Using Signalprints", in Proc. ACM WiSe'06, pp.43-52, Los Angeles, Sept. 2006.

[20]. L. Xiao, et al. "A Physical-Layer Technique to Enhance Authentication for Mobile Terminals," in Proc. IEEE ICC08, pp.1520-1524, Beijing, May 2008.

[21]. N. Patwari and S. K. Kasera, "Robust Location Distinction using Temporal Link Signatures", in Proc. ACM MobiCom07, pp.111-22, Montreal, Sept. 2007.

[22]. M.S. Bouassida, et al. "Sybil Nodes Detection Based on Received Signal Sybil Nodes Detection Based on Received Signal Strength Variations within VANET", International Journal of Network Security, vol.9, no.1, pp.22-33, July 2009.

[23]. P. Kalnis, et al. "Preventing Location-Based Identity Inference in Anonymous Spatial Queries", IEEE Trans. Knowledge and Data Engineering, vol.19, no.12, pp.1719-1733, Dec. 2007.

[24]. Yanchao Zhang, et al. "Securing Sensor Networks with Location-Based Keys", in Proc. IEEE WCNC'05, vol.4, pp.1909-1914, New Orleans, March 2005.

[25]. http://www.decawave.com/

[26]. http://spoonphone.com/en

[27]. C. Hennebert, et al. "Entropy Harvesting from Physical Sensors," in Proc. ACM WiSec'13, pp. 149-154, Budapest, 2013

[28]. C. Hennebert, et al. "The Entropy of Wireless Statistics," in Proc. EuCNC'14, Bologna, June 2014.

[29]. O. Gungor, et al. "Secret Key Generation via Localization and Mobility," http://arxiv.org/abs/1112.2793, v5, April 2014

[30]. S. A. Hoseinitabatabaei, A. Gluhak, R. Tafazolli, and W. Headley, "Design, Realization, and Evaluation of uDirect; An approach for Pervasive Observation of User Facing Direction on Mobile Phones," IEEE Trans. on Mobile Computing, vol. 99, no. 1536-1233, pp. 1–14, 2013.

[31]. T.Choudhury and A.Pentland,"Sensing and modeling human networks using the sociometer," Proc. Seventh IEEE International Symposium on Wearable Computers 2003, pp. 216–222, 2003.

[32]. G. Groh, A. Lehmann, J. Reimers, M. R. Friess, and L. Schwarz, "Detecting Social Situations from Interaction Geometry," 2010 IEEE Second International Conference on Social Computing, pp. 1–8, Aug. 2010.

[33]. J. Stehle, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Re´gis, B. Lina, and P. Vanhems, "High-resolution measurements of face-to-face contact patterns in a primary school." PloS one, vol. 6, no. 8, p. e23176, Jan. 2011.

[34]. A. Antoniou, E. Theodoridis, I. Chatzigiannakis, and G. Mylonas, "Monitoring physical space using mobile phones for inferring social and contextual interactions," Proc. 2011 IEEE

SENSORS, pp. 1616– 1619, Oct. 2011.

[35]. E. Miluzzo, N. D. Lane, K. Fodor, R. Peterson, H. Lu, M. Musolesi, S. B. Eisenman, X. Zheng, and A. T. Campbell, "Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application," in Proc. of the 6th ACM Conference on Embedded Network Sensor Systems, ser. SenSys '08, 2008, pp. 337–350.

[36]. N. Eagle and A. Pentland, "Social serendipity: mobilizing social soft- ware," IEEE Pervasive Computing, vol. 4, no. 2, pp. 28–34, 2005.

[37]. H. Lu, W. Pan, N. D. Lane, T. Choudhury, and A. T. Campbell, "Soundsense: Scalable sound sensing for people-centric applications on mobile phones," in Proc. of the 7th International Conference on Mobile Systems, Applications, and Services, ser. MobiSys '09, 2009, pp. 165–178.

[38]. Y. Wang and L. Cuthbert, "Bluetooth positioning using RSSI and triangulation methods," 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC), pp. 837–842, Jan. 2013.

[39]. A. Ghose, C. Bhaumik, and T. Chakravarty, "Blueeye: A system for proximity detection using bluetooth on mobile phones," in Proc. of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication, ser. UbiComp '13 Adjunct, 2013, pp. 1135–1142.

[40]. I. Carreras, A. Matic, P. Saar, and V. Osmani, "Comm2sense: Detecting proximity through smartphones," in Proc. Pervasive Computing and Commu- nications Workshops (PERCOM Workshops), 2012 IEEE International Conference on, 2012, pp. 253–258.

[41]. S. Liu, Y. Jiang, and A. Striegel, "Face-to-Face Proximity Estimation Using Bluetooth On Smartphones," IEEE Trans. on Mobile Computing, 2013.

[42]. W. Hu, G. Cao, S. V. Krishanamurthy, and P. Mohapatra, "Mobility- assisted energy-aware user contact detection in mobile social networks," Proc. 2013 IEEE 33rd International Conference on Distributed Computing Systems, pp. 155–164, 2013.

[43]. N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. Corner, "Virtual compass: Relative positioning to sense mobile social interactions," in Proc. of the 8th International Conference on Pervasive Computing, ser. Pervasive'10, 2010, pp. 1–21.

[44]. A. Matic, V. Osmani, A. Maxhuni, and O. Mayora, "Multi-Modal Mobile Sensing of Social Interactions," Proceedings of the 6th International Conference on Pervasive Computing Technologies for Healthcare, pp. 105–114, 2012.

[45]. P. Misra and P. Enge. Special issue on global positioning system. Proceedings of the IEEE, 87(1):3-15, 1999.

[46]. T. Garcia-Valverde, A. Garcia-Sola, H. Hagras, J. Dooley, V. Callaghan, and J. Botia. A fuzzy logic based system for indoor localisation using wifi in ambient intelligent environments. 2012.

[47]. L. Luoh. Zigbee-based intelligent indoor positioning system soft computing. Soft Computing, pages 1-14, 2013.

[48]. L. Ni, Y. Liu, Y. Lau, and A. Patil. Landmarc: indoor location sensing using active rfid. Wireless networks, 10(6):701-710, 2004.

[49]. N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell. A survey of mobile phone sensing. Communications Magazine, IEEE, 48(9):140-150, 2010.

[50]. N. Eagle and A. Pentland. Reality mining: sensing complex social systems. Personal and ubiquitous computing, 10(4):255-268, 2006.

[51]. J. Hightower and G. Borriello. Location systems for ubiquitous computing. Computer, 34(8):57-66, 2001.

[52]. S. A. Hoseini-Tabatabaei, A. Gluhak, and R. Tafazolli. A survey on smartphone-based systems for opportunistic user context recognition. ACM Computing Surveys (CSUR), 45(3):27, 2013.

[53]. H. Simon. Neural networks: a comprehensive foundation. Prentice Hall, 1999.

[54]. B. Li, T. Gallagher, A. G. Dempster, and C. Rizos. How feasible is the use of magnetic field alone for indoor positioning? In Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on, pages 1-9. IEEE, 2012.

[55]. D. W. Aha, D. Kibler, and M. K. Albert. Instance-based learning algorithms. Machine learning, 6(1):37-66, 1991.

[56]. N. Katzakis and M. Hori. Mobile phones as 3-dof controllers: A comparative study. In Dependable, Autonomic and Secure Computing, 2009. DASC'09. Eighth IEEE International Conference on, pages 345-349. IEEE, 2009.

[57]. B. Denis, J.-B. Pierrot, and C. Abou-Rjeily, "Joint Distributed Time Synchronization and Positioning in UWB Ad Hoc Networks Using TOA", IEEE Trans. on MTT, Special Issue on Ultra Wideband, Vol. 54, Is. 4, Part 2, pp. 1896-1911, April 2006

[58]. M. Pezzin and D. Lachartre, "A Low Power, Low Data Rate Impulse Radio Ultra Wideband Transceiver," in Proc. Future Network and Mobile Summit 2010, pp.1-10, Florence, June 2010.

[59]. D. Condeço et al., "Cooperative Location Algorithm Implementation and Evaluation", Deliverable D4.6 of the WHERE2 project, Dec. 2013

[60]. A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain," Image and Vision Computing, vol. 27, no. 12, pp. 1743–1759, Nov. 2009.

[61]. A. S. Pentland, "Automatic mapping and modeling of human networks," Physica A: Statistical Mechanics and its Applications, vol. 378, no. 1, pp. 59 – 67, 2007.

[62]. J. Hernández-Ramos, A. Jara, L. Marín, and A. Skarmeta, "Distributed Capability-based Access Control for the Internet of Things," Journal of Internet Services and Information Security (JISIS), vol. 3, no. 3/4, 2013

[63]. J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," Proc. IEEE SP'07, pp. 321–334, 2007.

[64]. http://www.usemp-project.eu/

[65]. A. Pérez, P. Larrañaga, I. Inza, "Bayesian classifiers based on kernel density estimation: Flexible classifiers," International Journal of Approximate Reasoning, Vol. 50, Is. 2, pp. 341-362, Feb. 2009

[66]. "IoT Communities and Identity Management", Deliverable D2.1 of the SocIoTal project, Aug. 2014.

[67]. I. Tunaru, B. Denis, B. Uguen, "Location-Based Pseudonyms for Identity Reinforcement in Wireless ad hoc Networks", to appear in Proc. IEEE VTC-Spring'15, Glasgow, May 2015

[68]. A. Fort, F. Keshmiri, G. Crusats, C. Craeye, and C. Oestges, "A body area propagation model derived from fundamental principles: Analytical analysis and comparison with measurements," Antennas and Propagation, IEEE Transactions on, vol. 58, no. 2, pp. 503–514, Feb 2010.

[69]. F. Ichikawa, J. Chipchase, and R. Grignani, "Where's the phone? a study of mobile phone location in public spaces," in Mobile Technology, Applications and Systems, 2005 2nd International Conference on, Nov 2005, pp. 1–8.

[70]. H. Liu and R. Setiono, "A probabilistic approach to feature selection - a filter solution," in 13th International Conference on Machine Learning, 1996, pp. 319–327.

[71]. R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artificial Intelligence, vol. 97, no. 1-2, pp. 273–324, 1997, special issue on relevance.

[72]. G. I. Webb, "Multiboosting: A technique for combining boosting and wagging," Machine Learning, vol. Vol.40, no. No.2, 2000.

[73]. R. Quinlan, C4.5: Programs for Machine Learning. San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[74]. E. T. Hall, "The hidden dimension", Doubleday, 1966.