Abstract—The advances of smartphone technology in recent years is leading to the uptake of a new class of Internet-based Indoor Navigation (IIN) services, which might soon diminish the need of Satellite-based technologies in urban environments. IIN services rely on geolocation databases that store indoor models comprising of floor-maps and points-of-interest along with wireless, light and magnetic signals used to localize users. Developing IIN services creates a new spectrum of information management challenges ranging from crowdsourcing indoor models, acquiring and fusing big-data velocity signals, localization algorithms, location privacy of custodians and others. In this work, we present the current landscape of academic and industrial IIN services using a multi-dimensional taxonomy of emerging topics in this domain, including location, crowdsourcing, privacy and modeling. We present the dimensions of our taxonomy through the lens of an open, modular, extensible and scalable IIN architecture, coined Anyplace, concluding with open challenges.

Keywords—Indoor, Location, Crowdsourcing, Privacy, Modeling.

1 INTRODUCTION

People in modern societies do 90% of their activities, business, commerce, entertainment and socializing indoors [30]. As all of these are increasingly aided by online services and as indoor spaces are becoming bigger and more complex, there is a growing need for cost-effective indoor localization, mapping, navigation and information services. People are nowadays equipped with omni-present mobile computing devices, which creates new opportunities for a variety of compelling applications in indoor spaces, such as, in-building guidance and navigation, inventory management, marketing and elderly support through Ambient and Assisted Living [11].

Additionally, there is a growing interest by retailers to enhance the shopping experience by offering on-the-spot coupons and by analyzing shopping behavior. Finally, the entertainment industry aims to design new games that exploit the actual environment of players as the playground [8].

Today’s technological market and gadget culture allow for the realization of such indoor services with the omni-presence of sensor-rich mobile devices in indoor environments. Mobile devices can measure a variety of signals, such as wireless, magnetic, sound and light, all relative to known locations in space (e.g., cell-towers, Wi-Fi Access Points (APs) [11] or beacons [2]). Organizing these signals in big-data geolocation databases, by possibly fusing the signals, allows to offer room-level (1-5 meters) or even sub-meter location accuracy [28]. The ideas presented herein secured the third overall position by Microsoft Research at IEEE IPSN’14 [28], with a localization accuracy of 1.96 meters. Combining indoor signals with detailed indoor context data, such as Points-Of-Interest (POI) collected with crowdsourcing [7] techniques by human custodians for monetary or ethical benefit, would frame what we would call an Internet-based Indoor Navigation (IIN) service.

In this article, we start out by rigorously classifying academic and industrial IIN services based on a multi-dimensional taxonomy we introduce, which includes localization, crowdsourcing, privacy and modeling. We also present an open IIN architecture, coined Anyplace [10] (http://anyplace.cs.ucy.ac.cy/), which has an open, modular, scalable and extensible architecture. The goal of Anyplace is to enable entities, such as individual users, companies or organizations to realize indoor applications using a scalable and multi-version information management approach.

2 TAXONOMY OF IIN SERVICES

In this section we provide a rigorous taxonomy of academic and industrial IIN services, using an accompanying Venn diagram shown in Fig. 1. The localization dimension is related to the requirement for dedicated equipment or not, which may heavily affect both deployment cost and accuracy. A large number of state-of-the-art geolocation systems rely on crowdsourcing [13] rather than cost-prohibitive data collection by professionals to address deployment cost, system scalability and maintenance bringing up, however, new research challenges. Privacy and confidentiality are critical for the wide adoption of indoor geolocation services because users have always been concerned about sharing their location data. While localization and privacy are key design factors, proper modeling of indoor spaces is equally important for the development of efficient IIN services and this dimension is not considered in existing classifications [11], [26], [13].

2.1 Localization

The core technology for IIN services is undoubtedly the type of hardware enabling the localization process. In the first dimension of our taxonomy we focus on whether new specialized hardware is needed to offer the location primitive, as opposed to either exploiting existing infrastructure (e.g., Wi-Fi network for the purpose of wireless connectivity) or no infrastructure whatsoever (e.g., inertial sensor systems). This classification is in line with recent evaluation efforts and field trials [28].
Infrastructure-based solutions require the deployment of additional dedicated equipment, including proprietary transmitters, beacons, antennas and cabling, for the provision of location signals. Industrial solutions in this field are termed Real-Time Locating Systems (RTLS) and current vendors specialize in specific markets for tracking assets using Active or Passive RFID, Bluetooth Low Energy (BLE) Smart beacons [2] (e.g., Apple’s iBeacon, Estimote.com context stickers or Quuppa.com sport beacons), Wi-Fi beacons (e.g., Aeroscout.com Active RFID Tags and Kontakt.io cloud/BLE beacon) or Ultra-Wide-Band (UWB) chips (e.g., DecaWave.com UWB transceiver offering a 10cm localization accuracy!) Assisted-GPS products (e.g., from Qualcomm) or multi-constellation solutions like CSR’s SiRFusion chipset that employs information from GPS, Galileo, GLONASS and Compass satellites, have the potential to improve availability in urban canyons and some indoor environments. However, to meet typical horizontal and vertical accuracy requirements indoors, both solutions fuse satellite measurements with other complementary source of information including cellular and/or Wi-Fi and/or multiple inertial sensors.

On the other hand, a prominent academic effort is the Epsilon system by Microsoft Research [24], which relies on LEDs that flash in millisecond intervals, so that only a smartphone camera can pick up the pulses (e.g., one commercial effort that uses this idea is Byтелight.com). Subsequently, a smartphone user can be localized on the intersection of circles defined by those distances (i.e., multi-lateration). The emerging Li-Fi standard extends the idea of LEDs into communications. Another academic effort is the ALPS system by Carnegie Mellon University [22], which employs ultrasound audio signals captured by the smartphone’s integrated microphone. Other older systems rely on wireless standards such as Bluetooth (BT) [4], which is too power hungry for the scenarios under discussion.

In theory, all aforementioned solutions have the potential to achieve sub-meter level accuracy at high deployment densities. However, these solutions raise scalability issues in case of large indoor spaces. For example Indoo.rs required 300 StickNFind beacons (i.e., estimated at 15,000 USD) to provide guidance to visually blind people at the 60,000 m² Terminal 2 of the San Francisco Intl. Airport, USA. Additionally, there are costs associated with maintaining the batteries of beacons and tuning their signal levels and transmission frequencies [2]. Generally, infrastructure-based approaches can be found in industries where the installation and maintenance costs can be compensated.

Indoor location revenues are forecast to reach 10 billion USD in 2020 according to ABIresearch [1]. Particular IIN services per market are:

- **Manufacturing**: Aeroscout.com, Ubisense.net;
- **Navigation**: PoleStar.eu, Indoo.rs;
- **Shopping Analytics**: Aislelabs.com, Walkbase.com;
- **Healthcare**: Versustech.com, Sonitor.com;
- **Museums**: Situate.io, Wifarer.com;
- **Internet-of-Things**: Estimote.com, Kontakt.io;
- **RTLS Hardware**: Zebra.com; DecaWave.com UWB.

Infrastructure-free solutions exploit location-dependent measurements from existing wireless communication infrastructure, such as Wi-Fi access points and cellular base stations. These solutions usually also leverage sensory data reported by Inertial Measurement Units (IMU), including accelerometers, gyroscopes and digital compasses (e.g., CyweeKIOS [25] or WiFiSLAM [9]). Such sensor modules are either integrated into modern consumer electronics, or attached externally on the human body, i.e., head, back, waist or foot mounted while processing occurs by low-power co-processor (e.g., M7-8 motion coprocessors on iPhones). Infrastructure-free localization often refers to solely IMU-based approaches that have been extensively studied in the context of Pedestrian Dead Reckoning (PDR) systems [32].

Infrastructure-free systems usually require only software modifications at the terminal side for the collection of location-dependent data [3]. Interestingly, recent field tests revealed that they are able to attain localization accuracy that is comparable or even better than costly infrastructure-based systems [28]. To reduce the data collection labor, several systems try to build a reliable indoor propagation model, which provides the Wi-Fi signal intensity (i.e., RSS) as a function of the distance from a transmitter (e.g., Wi-Fi AP) with known position [13]. Such a model is used to either generate the fingerprint database with significantly less effort or to calculate the distances between the device and neighboring transmitters to localize the user through multi-lateration.

In this category we can also classify a special type of RTLS that emerge by enterprise wireless LAN vendors. Companies such as Cisco CMX, Ericsson, Ekahau.com, Arubanetworks.com and Aerohive.com allow enterprises to manage their enterprise WLAN networks but also offer location tracking extensions to their services. Particularly, Wi-Fi routers with proprietary operating systems (e.g., using OpenWrt.org or Open-Mesh.com) can acquire the signal intensity of users that are moving inside a building with their Wi-Fi transceiver enabled. Acquiring these signals from several APs in a building
allows the creation of a user RSS fingerprint, which can subsequently be compared against a fingerprint database constructed a-forehand. This effectively allows the derivation of room-level localization accuracy down to a few meters (6-10 meters). Given that the APs can monitor the network MAC address of a mobile user, these approaches are fundamentally violating the location privacy of a user (even though the identity of the user is not exposed).

### 2.2 Crowdsourcing

Another way to classify indoor geolocation systems is based on whether the geolocation database has been populated by specialists or in a participatory fashion by non-experts.

**Non-participatory** systems usually employ paid professionals to undertake the data collection task, like in the *Ekahau* commercial system, or a small team of trained volunteers, like in the *KAILOS* academic project [12]. This involves one or more people visiting several locations that span the whole area of interest to collect a large volume of location-dependent data prior to positioning. This is not only laborious and time consuming, but it may also become cost prohibitive.

For example, it took 15 collectors about 2 weeks to collect point-by-point 200,000 Wi-Fi signal strength readings at 10,000 unique locations to cover the 450,000 m^2 COEX underground shopping mall area in S. Korea. Besides this, a measurement survey upon the *Ekahau* system installation can cost 10,000 USD for a large office building with no maintenance included [23]. To make things worse, localization data soon might become obsolete, e.g., Wi-Fi access points may be removed or relocated or new ones may be installed, which necessitates the collection of fresh data from time to time. For this reason such systems are more appropriate for big enterprises and the manufacturing industry (see RTLS).

In cases somebody only aims to map indoor models (POIs and maps) an indicative pricing given by Mazemaps in May 2015 was 0.17 USD per m^2 (i.e., 1,700 USD for a four story Departmental building).

**Participatory** systems like the traditional *Active Campus* [6] project introduced the concept of employing feedback from regular users to expand a core signal database created by trained contributors and keep maintenance cost low. Recently, crowdsourcing has emerged as a new paradigm to address maintenance and scalability issues [7]. Such solutions leverage localization data collected by common people while walking through corridors and rooms only clicking on the digital map to indicate turning and end points. The Intel *Place Lab* project [18] was among the first attempts to build geolocation systems that rely entirely on user collaboration, followed by recent systems including *Zee* [29], *FreeLoc* [31], *Mol´é* [23] and *Anyplace* [10] among others [13]. One example is *Anyplace*, in which 27 students at the University of Cyprus crowdsourced 12 buildings (36,000 m^2) each allocating a few hours per task in March 2015. The progress of crowdsourcing was available to the tasker through respective online heat-maps.

This approach is followed by key players like Google with their *Indoor Maps* project. One problem with participatory work is how to handle volume and noise. For example, Google suspended its Map Maker service in May 2015 because they could not validate the volume of user updates and because certain users were vandalizing Google Maps.

### 2.3 Privacy

A fundamental drawback of most network-based IIN services, is that the service can continuously “know” the location of a user while serving them. This problem has been referred to in the literature as *location privacy* [5]. On the other hand, transmission of complete geolocation databases to the user-carried device, to avoid compromising its privacy, means continuously transferring massive amounts of data through resource-limited wireless connections. Privacy-preserving mechanisms are discussed in [17] that also reviews studies of peoples’ attitudes about location privacy. We now focus on whether the IIN localization takes place on the terminal (i.e., offering privacy) or the network or cloud.

**Terminal-based (Privacy)** solutions calculate the location directly at the user-carried device, which hosts the localization algorithm, using sensor readings produced or downloaded locally. In this category, location privacy is easily preserved, network overhead is minimized and offline functionality is guaranteed as no information needs to be communicated between the IIN service and the user. The downside with these techniques is the increased battery depletion as the possibly complex localization computations take place on the device. Open systems in this category include our in-house *Airplane* [20] and Redpin.org.

**Network-based (No privacy)** approaches rely continuously on the IIN service, where the localization algorithm resides, and use location-dependent observations that are either monitored by the network infrastructure or collected by mobile devices in a terminal-assisted fashion. Such systems are mandatory in cases where terminals can not provide location-dependent measurements, e.g., signal strength observations from surrounding Wi-Fi access points are not available on Apple or Windows smartphones through operating system libraries, as opposed to Android smartphones. The open *SmartCampusAAU* IIN system features both terminal and network-based localization to support all three major mobile OS platforms [14].

Most major IIN services are currently network-based, including Google (Indoor), Trueposition (formerly Skyhook),Navizon.com, Insoft.com, Indoo.rs, and Wifarer.com, are all completely network or cloud-based and as such, can be considered to fundamentally compromise location privacy. Anyplace on the other hand, allows to choose between privacy and no privacy. In the former mode, users find their location either by downloading complete buildings or by downloading subsets of buildings through the IIN service without disclosing their location-context metadata [15].

### 2.4 Modeling

Unlike outdoor environments, indoor spaces are characterized by complex topologies and are composed of entities that are unique to indoor settings, such as multiple floors, rooms and hallways connected by doors, walls, stairs, escalators, and
elevators. To make things worse, doors may be one-directional (e.g., in security control in airports), while temporal variations may occur (e.g., a room may be temporarily available due to its opening hours, or conference hall might be partitioned into several smaller rooms to accommodate different events).

It is clear that the traditional Geometric (i.e., Euclidean) modeling approach is not appropriate. For instance, a location may not be directly accessible from another nearby location, even though their Euclidean distance is small because of a wall or floor. Symbolic modeling approaches address these limitations by using reference points (e.g., rooms) to establish a structure for distance computation. In general these approaches leverage a graph-based model [27], which allows direct usage of graph algorithms, e.g., shortest path, connectivity, traversals.

Most modern IIN Services (see Fig. 1) currently provide some kind of proprietary modeling extensions. There are some that specialize solely on mapping, rather than location itself, such as Micello.com, which has over 50,000 buildings worldwide and charging 49 USD per building per month to allow in-app integration; and Indoor.io, which focuses on amazing 3D maps (with location provided by Polestar.eu and Indooratlas.com). These services then deploy tile-servers that efficiently serve floor-tiles and POIs through API calls and with client side libraries (e.g., native libraries for popular smartphone OSes or javascript libraries). From the industry viewpoint, certain efforts rely on Autodesk’s Industry Foundation Classes (IFC) data model, which is open and registered by ISO, for the description of building and construction data. Using the given model, indoor spaces can derive structural semantics directly from Autocad files used by building architects.

2.5 Further Classifications

The proposed taxonomy is by no means exhaustive and in fact various performance evaluation criteria, including cost, complexity, robustness, scalability and commercial availability, could also be considered. We refer the interested reader to the following works that attempt to evaluate and compare indoor geolocation systems with respect to some or all of these criteria. We point out, however, that assessing system performance in terms of a specific criterion might not be a trivial task. For instance, monetary cost typically includes equipment, cabling, installation and setup costs, as well as maintenance and running costs. These are hard to quantify in the general case and attempting a simple categorization, e.g., as low, medium or high cost, might be misleading. Thus, these performance findings should be revisited depending on the application scenario and after prototype deployments move into real production deployments.

Further classifications can be adapted from related surveys and taxonomies. Authors in [11] classify indoor geolocation systems based on the signaling technology used to determine location and identify six main categories, namely infrared signals, ultrasound waves, radio frequency, electromagnetic waves, vision-based analysis and audible sound. In [26], indoor positioning systems are classified based on the measuring principle, i.e., proximity, angular, timing, or signal strength measurements, and the underlying positioning algorithms, including triangulation, trilateration and the increasingly popular fingerprinting approach. Both taxonomies are purely geared towards localization while we focus on IIN services. Recent surveys focus on Wi-Fi-based geolocation systems due to their increasing popularity and suggest technology-specific classifications based on the requirement for building map, explicit or implicit user participation and device heterogeneity among others [13].

3 The Anyplace IIN Service

Anyplace[10] is an infrastructure-free IIN service that leverages collaborative sensing and the availability of rich location-dependent data on smartphones to determine user location through local processing. Our platform consists of five main components, including the Server, the Architect, the Viewer, a datastore, and a client application running on Android smartphones acting as a Logger and a Navigator.

3.1 Localization in Anyplace

The combined Navigator and Logger is a designated tool for Android users, which can benefit from Wi-Fi fingerprinting [10], [28] available under this platform. The Navigator allows users to see their current location on top of the floorplan map and navigate between POIs inside the building, similarly to the Viewer (iOS, Android, Windows). The main difference, is that the Navigator offers superb accuracy. Our developments are available on Github under an MIT license.

The Navigator also uses the onboard smartphone sensors (i.e., accelerometer, gyroscope and digital compass), which are seamlessly integrated in our tracking module to smooth the Wi-Fi locations and enhance the navigation experience. We currently investigate the intelligent prefetching of Radiomaps [16] using historic analysis of trajectories. The Logger application enables users to record RSS readings from nearby Wi-Fi APs and upload them to our Server through a Web 2.0 API (in JSON). It is used by volunteers for contributing RSS data, and for crowdsourcing signal maps. In order to facilitate the collection of quality Radiomaps, we present a heat-map of previously collected fingerprints in the building.

3.2 Crowdsourcing in Anyplace

Anyplace features several modules to support crowdsourcing of location-dependent sensor readings collected on smartphones. Firstly, given that mobile devices are outfitted with diverse hardware sensors provided by a wide variety of vendors means that Wi-Fi measurements can greatly vary. Anyplace handles this by means of a differential fingerprinting module that outputs signal strength differences [21], instead of absolute values. Secondly, the outlier filtering module detects and rejects invalid user contributions to avoid the contamination of the Wi-Fi radiomap with erroneous signal strength data. This may occur accidentally, in case well-intentioned contributors click on the wrong part of the building to mark their true location while collecting data, or deliberately in case a malicious user aims to compromise the accuracy of the system.
3.3 Privacy in Anyplace

Anyplace offers a flexible privacy scheme, where a user has the option to localize by caching complete indoor models on the smartphone (thus obtaining absolute location privacy), or by intelligently downloading subsets of buildings through the IIN service without disclosing location-context metadata of users. A proposed Temporal Vector Map (TVM) algorithm [15] camouflages a user trajectory among k other users offering energy-efficiency, high performance in terms of retrieval time and network resource conservation, without hindering the provision of fine-grained location updates.

3.4 Modeling in Anyplace

The Architect Web app offers a feature-rich, user-friendly and account-based interface for managing indoor models in Anyplace. Particularly, after logging in a user can place the blueprint of a building on top of Google Maps with multi-floor support. Using the floor editor, the user can upload, scale and rotate the desired blueprints to fit them properly. The user can later add, annotate and geo-tag POIs inside the building and connect them to indicate feasible paths for enabling the delivery of navigation directions.

4 Future Challenges

In this section we enumerate a number of future challenges in the four dimensions addressed in this article.

Localization: A major future challenge is to fuse multi-modal location-dependent sensing data, coming from highly diverse, low-cost and error-prone smartphone sensors, in order to push indoor accuracy to the limit. Another challenge is the provision of seamless indoor/outdoor transition and transparent cross-floor movement. Finally, as indoor geolocation solutions become ubiquitous different systems will overlap and may compete for the provision of location information in some areas, which asks for intelligent switching from one solution/technology to another, based on availability, power consumption or accuracy requirements.

Crowdsourcing: This is a viable solution for the construction and maintenance of indoor signal maps, however it poses new challenges. When applied to global scale, it generates a continuous flow of user-contributed data that may quickly become overwhelming. This data explosion requires efficient big data processing architectures [19] for offline (e.g., hadoop-oriented) or online (e.g., spark-based) data processing and machine learning. To make things worse, sensor signals change dynamically over time, e.g., APs are removed and relocated, or new APs are installed, thus a signal map may soon become obsolete. An outdated radiomap may affect outlier detection and filtering capabilities of the geolocation system, resulting in fresh valid data being rejected.

Privacy: Location tracking by IIN services poses a serious imminent threat, which will have a much greater impact than other existing forms of location tracking, as it can occur at a very fine granularity in indoor spaces. Moreover, IIN are private enterprises that are less controlled, thus they might be tempted to exploit the “big” location data of their customers, by either selling it to advertising companies or by linking it to other sensitive data sources. Additionally, a user cannot know where IIN host and operate their data and whether these conform or not to latest legislative efforts and reforms. Finally, IIN are attractive targets for hackers, aiming to steal location data and carry out illegal acts (e.g., breaking into houses). Developing hybrid techniques, that on the one hand exploit the IIN utility, but on the other hand also offer controllable location privacy to the user, is certainly the way to go.

Modeling: The technology roadmap is towards indoor GIS integration where IndoorGML.net, GeoJson.org, or any other standard that may appear in the future, become fully interoperable. For instance, the ongoing European project i-locate.eu
aims to bridge IndoorGM and Open Street Maps (OSM) by building upon the lessons learned from the deprecated IndoorOSM tagging schema and extensions. Having the right modeling primitives will give rise to a variety of data management and query processing challenges in the future, such as effective in-building search and exploration. For example, a query “Guide me to all paintings in a museum that are of type Oil-on-Canvas”, could be answered effectively if indoor semantics were correctly modeled and fused with Knowledge-Bases such as Wikidata or YAGO.

5 CONCLUSION

This article summarizes the growing space of IIN services that aim to transform digital services in indoor and urban spaces. We provide a rigorous taxonomy that classifies many recent academic and industrial technologies and services, based on a rigorous multidimensional taxonomy. We present the dimensions of our taxonomy through the lens of an open, modular, extensible and scalable IIN architecture, coined Anyplace, concluding with open technical challenges in the field.

REFERENCES


6 ABOUT THE AUTHORS

Demetrios Zeinalipour-Yazti is an assistant professor at the Department of Computer Science at the University of Cyprus, Cyprus. Contact him at dzeina@cs.ucy.ac.cy.

Christos Laoudias is a post-doctoral researcher at the KIOS Research Center at the University of Cyprus, Cyprus. Contact him at laoudias.christos@ucy.ac.cy.

Kyriakos Georgiou is a graduate student at the Department of Computer Science at the University of Toronto, Canada. Contact him at kgeorgi10@cs.utoronto.ca.

Georgios Chatzimilioudis is a post-doctoral researcher in the Department of Computer Science, University of Cyprus, Cyprus. Contact him at gchatzim@cs.ucy.ac.cy.