Adaptive Bootstrapping for Crowdsourced Indoor Maps

Georgios Pipelidis, Christian Prehofer and Ilias Gerostathopoulos
Fakultät für Informatik, Technische Universtität München, Munich, Germany
{georgios.pipelidis, christian.prehofer, ilias.gerostathopoulos}@in.tum.de

Keywords: Indoor Mapping; Crowdsourcing; Bootstrapping Process

Abstract: Indoor mapping is an important and necessary enabler for many applications. However, indoor places and their services are very diverse. Furthermore, many technical approaches for indoor mapping exist. While there is fruitful research on combining some of these techniques, we show the need for flexible, customized bootstrapping for indoor maps. This includes mapping techniques but also intermediate services which enable data collection for improving maps and offering enhanced services. We illustrate examples of customizations of the process in a visual way and argue that the bootstrapping process needs to be adapted to specific buildings and end-user needs. This process-based view to indoor mapping leads to several research questions regarding the composition and intermediate steps in such process.

1 INTRODUCTION

Indoor mapping is an important enabler for many applications such as indoor navigation systems or for locating points of interest inside a building. This is a useful service even if indoor localization is not available. Together with indoor localization techniques, which have been an active area of research recently (Mautz, 2012), indoor mapping can help materialize the vision for ubiquitous indoor positioning system on a worldwide scale (Alzantot and Youssef, 2012).

There is considerable progress in the mapping of indoor places, and many diverse techniques have been proposed, ranging from robot-based (El-Hakim and Boulanger, 1999), vision-based (Gao et al., 2014), up to crowdsourced mapping (Alzantot and Youssef, 2012). However, most of the existing techniques are either expensive or difficult to apply, due to prone to error sensors and methods, and the variety of the building structures. It remains a challenge to provide cost-effective, easy-to-apply mapping techniques which can cover the large volume and variety of indoor places with their often unique characteristics and semantics.

Compared to outdoor maps, indoor mapping is more challenging for several reasons: *Indoor places are very diverse* in nature and many of them also change frequently; consider e.g. remodeling of floors or new shops in a shopping mall. Secondly, *indoor mapping techniques are very diverse* and range from manual with ad hoc tuning to crowdsourcing techniques. While manual techniques are often more reliable, the abundance of new personal devices with advanced sensors (e.g., motion sensors, cameras, gyroscopes, pedometers) also enable sophisticated crowdsourcing of indoor maps (Alzantot and Youssef, 2012). Third, the *services related to indoor mapping are also very diverse* in terms of end-user needs and technical assumptions. For instance, architects have different needs than pedestrians or fire fighters. Also, some services require localization, some only mapping, and some only user traces or landmark identification.

To emphasize the diversity of end-user needs and assumptions in the services related to indoor mapping, consider a hospital: the main service is finding doctors, patients, or equipment, assuming a well administered building with well defined tags for tracing and localization. Here, manually created maps can be used—a costly, yet worthy, investment for the hospital administration. On the other hand, in a shopping mall with diverse shop owners, diverse infrastructure and no central management of tags, users also aim to discover places, find other people and explore the map. Here, users may have time to contribute to crowdsourced map creation in exchange for some useful apps. Finally, in an automated factory, highly accurate indoor maps can be important in guiding robots, augmented reality and help avoiding accidents.

Following the above, in this paper we argue that there will be no single way for mapping indoor places, but rather *a diverse set of techniques and services will be used to build up maps and services for indoor locations in a customized way.* Some services may actually not even require proper maps, as in the case of a "take me to the exit" service for which only user traces can be sufficient. We also posit that we will move towards custom solutions for combining indoor mapping techniques in order to improve accuracy and enable a number of diverse services.

This position paper focuses on the combination of indoor mapping techniques and the services they enable. It specifically targets the problem of obtaining the critical mass of user data for self-starting crowdsourcing mapping techniques. In particular, we contribute by highlighting the need for a bootstrapping process that can be customized to the available techniques and building characteristics and by providing an example of such a process.

The rest of the paper is structured as follows. Section 1.1 overviews the most promising indoor mapping techniques. Section 2 provides an overview of our approach, while Section 2.1 exemplifies it on a specific bootstrapping process. Section 3 provides a short assessment of the current state of the art, while Section 4 puts forward a research roadmap and concludes by summarizing the key points.

1.1 MAPPING TECHNIQUES

We describe here the most prominent techniques for creating indoor maps.

Light Detection And Ranging (LiDAR). LIDAR uses lasers to measure the distance between objects inside a building (i.e., walls, floors, ceilings etc.) (El-Hakim and Boulanger, 1999). A LiDAR unit, often mounted on a robot or vehicle, scans the environment. The position of the unit is estimated by vSLAM (Karlsson et al., 2005). A point cloud is generated and by identifying contours (i.e. points of similar distance), a map can be extracted. Semantic annotations are usually manually made by expert surveyors.

Usage of existing architectural blueprints. If blueprints are encoded in formats such as Industry Foundation Classes (IFC) (iso, 2013) or Building Information Modeling (BIM) (iso, 2012), they contain the geometric information that can be readily used in indoor maps. However, such formats do not include topological nor semantic information. The last is usually added manually by expert surveyors, resulting into mapping data encoded into formats such as IndoorGML (Ind, 2016). Approaches for automatic derivation of topological relations (e.g., adjacency and connectivity of rooms) from IFC models have also been suggested (Liu et al., 2014).

Structure from motion. In this technique, a 3D structure of a building can be extracted from a camera (Gao et al., 2014) by capturing many images of an indoor place and translating them into a single 3D view. To do this, the camera's internal and external parameters, e.g. lens-generated distortion, translation and rotation matrix have to be known or be retrievable from common features of the captured images.

Depth sensors. In this technique, a typical setting is to have an infrared projector that projects a unique pattern. An infrared sensor, whose relative distance to the projector and rotation are known, recognizes this pattern. A depth map is constructed by analyzing the unique pattern of infrared light markers by triangulating the distance between the sensor, projector and the object. Finally, a 3D point cloud is extracted from stereoscopic view algorithms, from which a map can be generated (Henry et al., 2012).

Smart phone 3D modeling tools. In this technique, specialized smart phone apps enable users construct components of a building (Eaglin et al., 2013). After initial versions of the maps have been created, other users can enhance the maps or vote on their accuracy and completeness.

Activity-based map generation. An indoor map can be transparently and autonomously generated based on activity recognition of users (Alzantot and Youssef, 2012). This technique works as follows: After extracting steps of users by their x and y coordinates or by a series of trajectories, a point cloud can be extracted. A map of the indoor place can be created by fusing data from different users and identifying places with common patterns. For example, places where users performing the same activity (i.e., stairs) can be identified.

2 ADAPTIVE BOOTSTRAPPING

In this section, we outline our envisioned approach towards indoor mapping, based on the following observations on the present and future research and development in indoor mapping:

- *Techniques need to be combined.* There are many indoor mapping techniques which differ in terms of complexity, required resources, and output. For instance, if one wants to use LiDAR, a localization technique has to be in place, and also sophisticated laser equipment has to be available. Activity-based map generation, on the other side, does not make any major assumptions in terms of equipment; however, it assumes a plethora of data. We argue that a combination of different techniques will be used to create or maintain indoor maps that are both cost-effective and accurate.

- *Bootstrapping is needed for crowdsourcing.* As discussed, we posit there will be no "single-shot" solution towards indoor mapping; combined solutions, as shown below, will also involve crowdsourcing. Therefore an incremental, stepwise bootstrapping will be needed to obtain user data.
- *No single bootstrapping process.* We believe that the diversity of buildings, mapping techniques, as well as services will lead to individual and custom processes for such bootstrapping. The processes will be adapted to end-user needs, available infrastructure, available budget, and other factors.

A number of services with different characteristics, users, and assumptions on crowdsourcing effort can be supported by our approach, e.g.:

Wellness monitoring. This is a family of emerging services that provide feedback to users based on their activities during the day. For example, services that can track the number of steps that a user did during a day can be used for identifying the distance traveled by the user.

Card swiping. This service may substitute the Magnetic stripe cards with smart phone build-in NFC chips. In combination with other sensor data, it can be used to generate a general model for identifying outdoor-indoor transitions and vice versa.

"Take me to the exit". This service can work as a digital Ariadne's thread, where users will be able to find their way back to the entrance of indoor places by following their own captured route in reverse. User traces collected from this service can be used for generating a point cloud.

Instruction-based navigation. This service can provide basic instructions on how to visit an office or a classroom in the form of instructions such as "Enter from the north entrance, walk straight for 10 secs, then turn right, walk up the stairs and enter the door on the right".

Elderly monitoring. This service can be used to identify accidents involving elderly or people with special needs in real time by detecting problems in mobility or patterns that correspond to sudden falls. Data from such service can be used for semantically enhancing indoor maps, via adding the use of a room.

Dynamic meeting scheduler. This service can use the (indoor) user position in order to propose meeting locations that fit the participants' locations. Data from this service can be used for labeling indoor spaces.

It is clear that the services related to indoor mapping are rather diverse, and make different assumptions regarding the maturity and completeness of the supporting indoor mapping systems. For instance, wellness monitoring does not assume any complete mapping or localization system (even though the data captured from such services can actually allow for activity-based mapping techniques). Also, "take me to the exit" does not assume the existence of a complete navigable map, but only of a single well-defined route from a *single* user.

An important observation is that services with rudimentary assumptions in terms of indoor mapping can act as catalysts for gaining the critical mass of user data that can enable services with more advanced mapping needs. For instance, in a hospital building, the target service might be full-blown indoor navigation, whereas intermediate services might be call forwarding for medical personnel, room-based localization of equipment, elderly monitoring, and others. Potential users are the medical personnel, patients, and visitors. In contrast, consider a university campus building: the target service can be the same as in the hospital case, but now intermediate services could be room finders, "take me to the exit", wellness monitoring, etc., whereas potential users are now students and academic employees. Finally, in the case of a subway station, a promising intermediate service is, e.g., location-aware ticketing.

In the following, we are providing a way to model such bootstrapping processes. Our modeling technique is based on the fact that each indoor mapping technique can be broken down to a number of tasks with inputs and outputs. The input of the initial task indicates the technique's assumptions. As a result, a bootstrapping process can be represented as a graph of tasks. We present an example of this in the next section.

2.1 Bootstrapping Example

This section introduces an example of a bootstrapping process for a university campus building. To illustrate the bootstrapping process, we use a data-flow-like diagram depicted in Figure 1.

In this diagram, circular nodes correspond to artifacts. Each artifact enables the creation of one or more services. For example, `Distance Traveled` (`e`) can enable a service such as wellness monitoring, since the walked distance is directly related with exercising. Inputs and outputs of artifacts are visually presented as solid enumerated arrows which indicate data

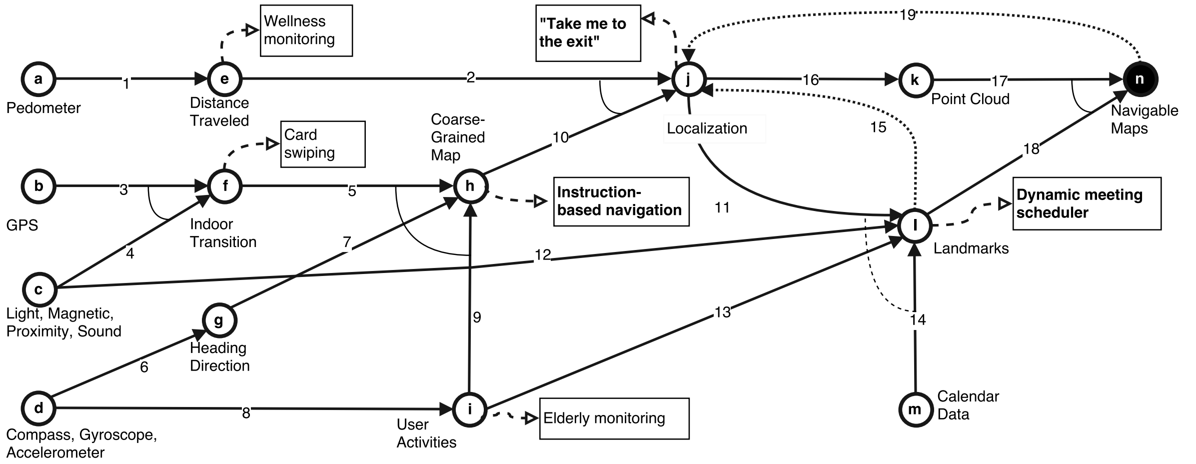

Figure 1: Customized bootstrapping process for a university campus building. Circular nodes are artifacts, arrows are tasks with inputs and outputs, rectangles are intermediate services (services in bold are described in the text).

flow. For example, the input of `Indoor Transition` (`f`) is GPS signal (`3`) and IMU (`4`) data (i.e. ambient light, magnetic field, proximity and sound). By reasoning on these input data, similar to (Zhou et al., 2012), the output is the locations of entrances (`5`). In case of more than one input, a solid line connecting them implies conjunction (e.g. lines `5`, `9` and `7`); a dashed line implies disjunction (e.g. `11`, `12`, `13`, `14`). Finally, dotted connections imply additional inputs which can improve the data quality (e.g. `15`).

An artifact can be connected to a number of intermediate services. A service is represented by a rectangle and implies a set of software functionalities which can be a user-facing application. Finally, the target artifact is represented as a filled circular node (e.g. `n`).

Figure 1 presents a set of possible bootstrapping options. One would start at one or more of the nodes on the left, e.g. assuming devices with GPS (`b`) or compass/gyroscope and accelerometer (`d`). Informally speaking, we can then proceed to some of the connected nodes (e.g. `f` or `g`), based on user data generated from operating services possible at this point. Based on the new data, we can proceed with further steps in this graph.

As depicted in Figure 1, the entire bootstrapping process could emerge through existing services, such as wellness monitoring or card swiping. Of course, alternative paths are also available. For example the `Coarse-Grained Map` step could be skipped; similarly, `User activities` might not be needed if semantically-rich calendar data are available.

In our example, the target service is to enable indoor navigation based on dynamically created maps that capture the geometry, topology and semantics of the building. The above information needs to be integrated in a data model, e.g. by using and extending the IndoorGML standard (Ind, 2016). IndoorGML provides the constructs to denote subdivisions of indoor places (i.e. rooms), spaces that connect two indoor places (e.g., inner doors), spaces that connect indoor places to outdoor ones (e.g., entrance doors), spaces acting as passages between indoor places (e.g., corridors, stairs), and other important properties.

There are a number of intermediate services among the ones described in the beginning of this Section. We describe here the indoor mapping techniques and associated artifacts they rely upon:

Instruction-based navigation. To provide this service, a `Coarse-Grained Map` is needed. This is a model that includes the elements essential for routing, such as corridors, stairs, doors, and entrances. This is the outcome of merging three other artifacts: `Indoor Transition`, `Heading Direction` and `User Activities` (tasks `5, 7, 9`). The first one is derived by using GPS data (task `3`) and fusing them with other mobile sensor data such as light, magnetic, and proximity data (task `4`). The intuition is that the sensors' behavior changes during the outdoor-indoor transition, where the GPS uncertainty and the WiFi received signal strength are both increasing.

`Heading Direction` can be derived via machine learning algorithms (embodied in task `6`) that work on compass, gyroscope and accelerometer data. The intuition is, if a phone's pose is identified, it can be used to extract the user's local direction (i.e. in the phone's coordinate system) via monitoring the acceleration changes due to the gait movement, then relate this direction to a global system using the compass.

Finally, `User Activities` can be derived from the same data using machine learning techniques with high accuracy (task `8`), since moving and stationary activities can be detected from disturbances in the acceleration sensor, while movements on the vertical space can be detected from disturbances in the barometric sensor.

Dynamic Meeting Scheduler. This service is based on the `Landmarks` artifact. Landmarks are distinctive locations in a building. They are either locations where users consistently perform the same activity (e.g., stairs)—contributed by the `User Activities` (task `13`)—or locations with distinct characteristics of a measured quantity (e.g., WiFi RSS, geomagnetism, sound, light)—contributed by the `Light`, `Magnetic`, `Proximity`, `Sound` (task `12`). In both cases, landmarks need to be localized in a building—hence the dependence on `Localization` (task `11`). Landmarks can also be derived from `Calendar Data` (task `14`) via semantics (e.g., meeting room name).

"Take me to the exit". In our example, we assume that there is no localization infrastructure in place. As a result, we would need to resort to pedestrian dead reckoning techniques (Kourogi and Kurata, 2014). Pedestrian dead reckoning is based on approximating the position of a user by measuring the distance traveled when walking towards a direction from a known point. This explains why `Localization` depends on the `Distance Traveled` (task `2`) and the `Coarse-Grained Map` (task `10`). The former is derived directly from pedometer data (task `1`). The latter contains information regarding the heading direction (task `7`) and the indoor transition points (task `5`). These points are the initial *known points* in the dead reckoning algorithm. `Localization` can also depend on `Landmarks` for re-calibrating the algorithm (restarting the error) in distinct locations (task `15`).

Finally, `Localization` provides input for the creation of `Point Cloud` (task `16`) using existing techniques, and subsequently of `Navigable Maps` (task `17`). `Navigable Maps` are also enhanced by the identified `Landmarks` (task `18`). In particular, activity-related landmarks can be a rich source of semantic annotation for maps (e.g., places where people sit together for long time can be labeled as meeting rooms). At the same time, `Navigable Maps` can enhance `Localization` by error recalibration on the basis of non-navigable places (task `19`). This can be achieved either by relating user traces to sets of possible routes or via uniquely identified locations (e.g. stairs), in which case the context of users (e.g. "climbing stairs") can be used for re-positioning them.

It is important to note that the example bootstrapping process illustrates a cost-effective solution without dedicated equipment and expensive manual work. As an alternative, consider hiring an indoor localization company, for performing tasks `1` and `2` in our example—this would have led to a different customization of the same bootstrapping process.

3 RELATED WORK

To our understanding, there is no prior work on systematic bootstrapping of indoor maps. There are several works which integrate different intermediate techniques, which we list below.

Heading direction. (Roy et al., 2014) detect the discreate signal vibration when the heel strikes the ground during a gait circle. Then they use this data point as a reference and scan the signal to identify the dominant body's movement partition from the entire signal segment. Finally, they translate the walking direction to the global magnetic system. However, their framework is highly dependent on the terrain as well as on user behavior.

Indoor-Outdoor transition. (Zhou et al., 2012) do not only use the drop of GPS accuracy as an indication of the I/O transition, but also use light censors, cell tower signals, and magnetic field sensors. The acceleration and proximity sensor time series are fused for identifying the I/O transition.

Activity Recognition. (Nguyen et al., 2015) use a Support Vector Machine classifier to distinguish among moving activities such as walking, running, and ascending and descending stairs and improve existing position systems. Their observation is that the step length varies when a user is walking, running or climbing stairs. Their approach is argued to work in various phone poses. However, their approach uses a large amount of features, which can result in high computational demands.

4 DISCUSSION AND OUTLOOK

Following the diversity of indoor places, techniques and services, we have outlined our position for an adaptive bootstrapping process. This includes mapping techniques but also intermediate services which enable data collection for improving maps and offering enhanced services. We have illustrated examples of customizations of the process in a visual way and argue that the bootstrapping

Our view integrates many existing mapping techniques as well as services and also assumes considerable progress in each of these disciplines. As we focus more on how the different processes for mapping can be integrated, our vision is orthogonal to research roadmaps of specific techniques.

Our new bootstrapping approach also gives rise to the several challenges:

Bootstrapping processes. We need research to understand and model bootstrapping processes, similar to our example, in order to obtain a more complete picture of the techniques and services that are available. Also, most of the services described in Section 2.1 are open challenges mainly due to the inherent complexity of indoor localization: existing sensors (both in phones and specialized devices) fail to effectively propagate a discrete signal patterns in indoor space, making simple triangulation-based techniques infeasible. Additionally, robust heading direction identification independent of the phone's pose remains an open challenge (Zhou et al., 2012).

Intermediate targets/artifacts. We need to understand what can be useful intermediate targets/artifacts, which are both feasible w.r.t mapping techniques and also enable useful services. Moreover, protocols need to be emerged to enable information exchange through APIs between the different services. Importantly, we need to manage the uncertainty inherent to both sensor reading and human users, filter out outliers, and in general work with noisy data. Trust models to manage ambiguous information extracted from multiple users need to be emerged. Existing indoor data models have to be enhanced in order to cope with such incomplete, ambiguous or inaccurate models.

Process customization. We need research to understand when and how to apply different bootstrapping processes to specific buildings. This can also lead to easier or automatic customization of bootstrapping to specific classes of buildings.

ACKNOWLEDGMENTS

This work is part of the TUM Living Lab Connected Mobility project and has been funded by the Bayerisches Staatsministerium für Wirtschaft und Medien, Energie und Technologie.

REFERENCES

(2012). ISO/TS 12911:2012 - Framework for building information modelling (BIM) guidance.

(2013). ISO 16739:2013 - Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.

(2016). OGC IndoorGML version 1.0.2. `http://www.opengeospatial.org/standards/indoorgml`.

Alzantot, M. and Youssef, M. (2012). CrowdInside: Automatic Construction of Indoor Floorplans. In *SIGSPATIAL '12*, pages 99–108. ACM.

Eaglin, T., Subramanian, K., and Payton, J. (2013). 3D modeling by the masses: A mobile app for modeling buildings. In *Proc. of PERCOM '13 Workshops*, pages 315–317. IEEE.

El-Hakim, S. F. and Boulanger, P. (1999). Mobile system for indoor 3-d mapping and creating virtual environments. US Patent 6,009,359.

Gao, R., Zhao, M., Ye, T., Ye, F., Wang, Y., Bian, K., Wang, T., and Li, X. (2014). Jigsaw: indoor floor plan reconstruction via mobile crowdsensing. In *Proc. of MobiCom '14*, pages 249–260. ACM.

Henry, P., Krainin, M., Herbst, E., Ren, X., and Fox, D. (2012). RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments. *Int. J. Robot. Res.*, 31(5):647–663.

Karlsson, N., Di Bernardo, E., Ostrowski, J., Goncalves, L., Pirjanian, P., and Munich, M. E. (2005). The vslam algorithm for robust localization and mapping. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 24–29. IEEE.

Kourogi, M. and Kurata, T. (2014). A method of pedestrian dead reckoning for smartphones using frequency domain analysis on patterns of acceleration and angular velocity. In *Proc. of PLANS '14*, pages 164–168. IEEE.

Liu, H., Shi, R., Zhu, L., and Jing, C. (2014). Conversion of model file information from IFC to GML. In *IGARSS'14*, pages 3133–3136. IEEE.

Mautz, R. (2012). *Indoor positioning technologies*. ETH Zurich, Department of Civil, Environmental and Geomatic Engineering.

Nguyen, P. et al. (2015). User-friendly activity recognition using SVM classifier and informative features. In *IPIN'15*, pages 1–8.

Roy, N., Wang, H., and Roy Choudhury, R. (2014). I am a smartphone and i can tell my user's walking direction. pages 329–342. ACM Press.

Zhou, P., Zheng, Y., Li, Z., Li, M., and Shen, G. (2012). Iodetector: A generic service for indoor outdoor detection. In *SenSys '12*, SenSys '12, pages 113–126. ACM.