A Novel Framework for Personalized and Context-aware Indoor Navigation Systems

Attila Török and Tamás Helfenbein
Institute for Applied Telecommunication Technologies (BAY-IKTI)
Bay Zoltán Nonprofit Ltd. for Applied Research
Budapest, Hungary
Email: {attila.torok, tamas.helfenbein}@bayzoltan.hu

***Abstract*—The recent indoor localization techniques use inertial sensors for position estimations in order to obtain a certain degree of freedom from RF solutions. Unfortunately, this dependency cannot be completely eliminated due to the cumulative errors introduced in the localization process; thus, RF or visual reference points are still necessary. In this paper we propose a novel approach for architectural design of indoor localization and navigation services by introducing a context-aware and extendable system framework. We exploit the ability to recognize certain human motion patterns and by using a scenario specific navigation language, for guiding the localization and position refinement process, we will be able to control the navigation on a much finer level. Therefore, in our system the reference points become needless or for scenarios with topological black holes at least the refinement process is automated, the user can be omitted from finding these points.**

***Index Terms*—indoor navigation; context-awareness; location based services; pedestrian localization.**

I. Introduction

With the advent of smart phones Location Based Applications (LBA) [1] witness an ever increasing popularity. While commercial services mainly focus on outdoor use cases indoor LBAs suffer a relative backlog, although at first sight all the necessary building blocks [2] are available. Besides the lack of common standards an even more stressful reason can be attributed to the scenario specific nature and sensibility to infrastructural changes of indoor localization techniques. Existing solutions require special effort to build detailed RF maps or propagation models and these pre-deployment steps must be repeated in case of variations (topology, transmission power) in system configuration. Therefore, the research community has started to focus on indoor positioning techniques where pre-deployment efforts are not necessary [3] or where the localization is based on minimal infrastructure [4]. To obtain a certain degree of freedom these systems leverage the technological advancements in mobile devices, such as the inclusion of accelerometer, compass or magnetometer sensors [5] [6]. However, in the current proposals the dependence from some kind of reference points (RF, visual information or acoustic beacons) cannot be completely eliminated, since due to the nature of the inertial sensors by distancing from the last known reference position cumulative errors will be introduced in the location estimation process.

In recent indoor navigation systems, besides the afore mentioned inertial sensor fusion (called dead-reckoning (DR)) techniques, camera phones can also support localization and navigation. These approaches [7] [8] [9] use well placed visual markers (e.g., 'YOU-ARE-HERE' (YAH) maps, QR codes) in order to provide reference points for DR drift cancellation. To further improve navigation experience besides the traditional map based solutions augmented reality (AR) interfaces are also applied. The traditional AR interfaces usually require continuous localization of the user, while in newer ones constraint diminution is achieved by using sparse localization techniques. Unfortunately, these solutions still require reference points [7] [9], constant user interaction/supervision [10] or an occasional manual reset of the accumulated location error [8].

A common problem with current indoor LBA solutions is the moderate effort dedicated to consider and deeper explore the dimensions of contextual relationships in the localization and navigation process, consequently the different requirements arose from personalization (user preferences/capabilities), scenario peculiarities (topological/service types) and their relationship to positioning/route guidance is not integrally handled. Different users will have different capabilities and requirements regarding the navigation procedure. For example, active participation in the process (navigation or interaction through AR interfaces) should be avoided, since it can distract the user, causing confusion/accidents. Instead, a proper voice guidance based navigation service shall be used. Considering scenarios, in certain premises the placement of any kind of reference points is beyond possibility due to legal-, investment issues, or their usability is just simply questionable. Also different levels of quality of service will be required for positioning and route guidance in miscellaneous scenarios.

In this paper, we propose a novel approach for architectural design of indoor localization and navigation services, aimed to provide an extendable framework for personalized and context-aware indoor LBAs. Our goal is to provide a navigation system, which requires no RF infrastructure and where the user interaction for finding the reference points is minimized or at least is automatically triggered by the scenario, the navigation process itself. This requires the employing of human movement behavior analysis, the introduction of a special navigation language, which controls the localization and position refinement process, and the design of a novel architectural framework to empower and piece together the building blocks of the system.

 ISBN: 978-1-61208-229-5

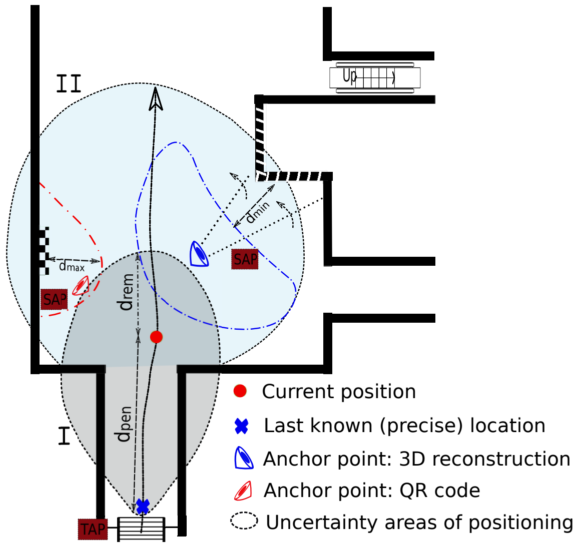

Fig. 1. Problem definition of indoor positioning

The paper is organized as follows. We present the identified indoor localization problems in Section 2, followed by the description of the system architecture. Section 4 presents the position refinement using cameras and we conclude the paper in Section 5.

II. Problem Statement

Navigation systems based only on inertial sensors are corrupted by cumulative errors. As this error grows, the positioning inaccuracy can cause local disorder of the navigation service. Refinement of the position estimations is done by fusing the results with absolute position measurements derived from the reference points of an additional localization system.

Since in our system we want to keep away from using reference points, which require mounting of special infrastructures (e.g., WiFi access points) we have to look for new ways to provide absolute positions for localization error cancellation. Considering, for example a floor map of a subway system we can identify certain building blocks like underpass areas, halls and platforms usually connected with tunnels, stairs and escalators. By traversing trough such places we will generate specific behavior patterns, such as walking, climbing stairs, making turns, taking elevators/escalators, etc. These actions can be considered as special events and by correlating the recognized events with the estimated movement pattern and the topology of the floor plan we will be able to refine the position estimation of the users. These topology specific points used for localization error correction we call Topological Anchor Points (TAPs). Therefore, we extend the functionality of inertial sensors to detect human motion related patterns by using real-time feature detection algorithms at the mobile side. In the current localization systems the accelerometer sensors are mainly used as a digital step counters [4] [5], the recognition of human activities is typically used in Assisted Living applications to detect daily activities of people (e.g., [11]) in health-care services.

Despite the introduction of this concept in certain scenarios there will be no specific topological points, which can generate particular, well interpretable events or the distance between two consecutive TAPs will be too long. In such cases (called topological black holes) the growth of the localization error (uncertainty area nr. II on Figure 1) cannot be suppressed efficiently, this leads to inaccuracy of positioning and to possible navigation problems (missing the second exit with the escalator). To cope with similar situations we also allow to create Soft Anchor Points (SAPs), which are not inherently topology specific and where the special events triggering error cancellation are derived by using additional techniques (e.g., by identifying visual markers like QR codes or 3D positioning via cameras). These SAPs will be defined by the users/service administrators, their usage requires some sort of user interaction, similarly to the visual reference points used in [7] [8] [9]. However, in our system the placement of SAPs is not hard-coded in the system, we do not expect the existence of fixed visual reference points (e.g., YAH maps), since in many cases these are not optimally placed, considering the navigation service needs. The possibility to dynamically place the SAPs can also facilitate the introduction of new location refinement methods, such as computer vision techniques. Since the SAPs are placed to unknown locations we have to notify somehow the users for triggering the refinement process.

To analyze the problem of SAP placement and refinement triggering we present the scenario of Figure 1, where we can observe a certain correlation between the favored placement of SAPs and the movement pattern of the user. As we can see, the last known precise user location is triggered by a TAP (stairs) and from this point the positioning error grows (uncertainty area) as the user enters into the hall. Despite the increasing localization error, until the user overpasses the uncertainty area nr. I (by traversing through the tunnel and entering the hall), there is no reason to trigger the refinement process, since no useful information can be provided for the navigation process. The growth of the localization error can be handled by DR and a context specific mechanism, which will keep the navigator on track, not letting to assume any unnatural events (e.g., crossing the tunnel walls), since is physically impossible to act differently due to the scenario's constraints. As the uncertainty area grows beyond a certain, well defined level, and the user is possibly situated in a more complex scenario (hall/underpass with many exits) the refinement process has to be triggered. This point of action will be calculated based on the floor plan, the nature of the inertial localization algorithms (estimated positioning errors) and the user preferences (e.g., do not use camera). Choosing a specific method for SAP creation will also affect the quality of the navigation process. For example, by using computer vision the area from where correct localization can be effectuated is larger than using QR codes, and it is also easier for the user to perceive and find the right spot (zones around SAPs). The context specific evaluation will also let us to provide an optimal placement for the SAPs, considering the floor plan, the navigation service requirements and the existence of scenario specific TAPs.

 ISBN: 978-1-61208-229-5

III. System Architecture

In order to preserve the flexibility of the framework, we propose the system architecture presented on Figure 2. One key design concept is the separation between the service specific and sensor related data plans. Service specific data (e.g., map tiles, route queries/responses, navigation instructions) are exchanged through the communication channels between the mobile and the respective service. This channel is also used to provide a navigation specific script for the mobile application, which we call Navigation Markup Language (NML).

To assure extendibility, all the information involved in the localization process is considered as sensor data and is exchanged through a Data Gathering Server (DGS); thus, we will be able to provide location related information for other services, too. The output of some sensors (usually accelerometer and magnetometer data) is processed at the mobile side by defining proper signal processing and classification rule sets, while other sensor's data (e.g., GPS) is collected as a result of a simpler query. The first kind of sensors (called Virtual Sensors (ViSe)) will let us to specify the feature detection algorithms used for human motion recognition, giving the ability to derive information for the localization algorithms (e.g., distance calculation based on step counter, DR) and also for the error correction algorithm using TAPs (e.g., by producing events in case of stair detection). The sensor data involved in localization and navigation is shared through the ViSe Routing Nodes, using a publish/subscribe communication graph. The capabilities, the ViSe defined on a specific mobile node, are published and shared through a commonly accessible Knowledge Base (KB), from where all the newly introduced services can acquire information (ViSe discovery). ViSe data and control information are separated using content based routing in DGS (by default the KB gets only control data), in order to facilitate the scalability and extendibility of the system. Thus, the transport technology used between the mobile and the service (e.g., XMPP) is detached from the service discovery, the method of accessing the KB (web service). Inter-service communication and eventing is also better supported, since the service advertisements can be channeled into the KB. The control of mobile sensors, related to the requirements of a specific service, is done by using the Service Control Nodes, whose membership has to be managed by the service or the KB itself.

In our system the Navigation Markup Language (NML) is used to control the localization and to provide context specific information besides the usual navigation related information (floor plan, navigation instructions). The NML is generated during the route planning phase, by analyzing the scenario itself (topology of the indoor environment, user preferences, etc). From the topology graph derived from the floor plan and the planned route the affected TAPs will be determined, acting as error cancellation points. The ViSes on the mobile related to the specified TAP recognition (providers of the respective movement patterns) will be asked for event reporting in form of queries. Finding the corresponding SAPs along a planned

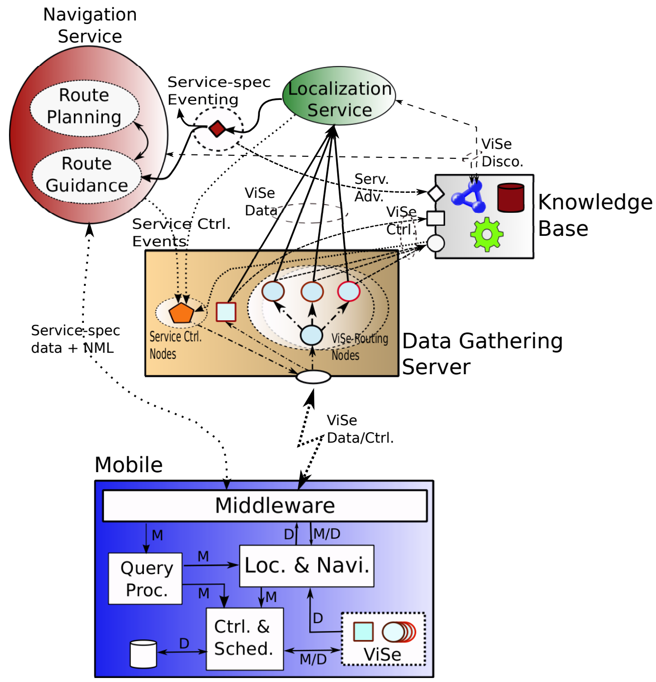

Fig. 2. System architecture of the navigation system

route and triggering the correction procedure is related to the definition of uncertainty areas between two consecutive TAPs (off-line calculation and configuration/setup). This task can be also formulated as an NML query and its result is triggered by an estimation procedure: when the distance between the last know precise location and the estimated actual position exceeds the threshold defined by d_{pen} + d_{rem} (see Figure 1). The refinement in the localization process and the possible recovery of the navigation are calculated based on the reported ViSe data on the server side.

IV. Position Refinement Using Cameras

Cameras can be also used as absolute position estimators. Researchers are trying to develop mobile resource effective methods using artificial landmarks (e.g., painted signs [9]) or statistical scene modeling. Unfortunately, these methods are sensitive to the (financial, aesthetic and judiciary) cost of landmarks and the dynamics of the modeled scene. Accordingly, in our case both pre-installed QR codes and 3D reconstruction based positioning will be used as checkpoints. We extract image key-points and calculate 3D coordinates using at least two digital images of the same checkpoint and its surroundings. We use two types of input: normal and panoramic images [12]. If two normal images are used as input, pose (position and orientation) of the camera needs to be recorded. Hence, this method is used by system installers. In case of panoramic images, the pose of the camera can be estimated relative to the reconstruction coordinate system. To correctly align the reconstructed model with the predefined spatial one, arrangement of branches, stairs or exits can be used. Thus, reduction of 3D point cloud can be used by slicing

Copyright (c) IARIA, 2012. ISBN: 978-1-61208-229-5

in the range of typical camera heights. From input images (normal and panoramic) distinctive invariant image features are extracted. Feature vectors are matched across images to triangulate key-point positions. Reconstruction tasks can be done on server or mobile side. During positioning, along the route, the feature vectors and 3D coordinates of a checkpoint are downloaded to the mobile node. Restricted search region of feature extraction and matching is used with respect to the possible camera pose derived from previous pose and inertial sensor data, to reduce computing costs on the mobile device. In order to provide fast localization, we decided to use the (so far) unexploited parallel processing capabilities of mobile GPUs. We are working on models and methods to estimate optimal QR code placements and reconstruction camera poses according to the spatial constraints of the scene, computational and geometric constraints of methods, and inaccuracy models of sensors.

V. Conclusion and Future Work

In this paper, we proposed a novel approach for architectural design of indoor localization and navigation services. Our goal is achieved by employing human behavior analysis to recognize certain movement patterns, such as walking, taking stairs, using elevators, etc. These events are used to define topology specific points, which recognized during the navigation process can suppress the localization error. For complex scenarios with topological black holes (where the distance between two consecutive TAPs is too large) we introduced the concept of SAPs, whose optimal placement can be calculated off-line and their finding can be triggered automatically by the navigation language (NML). The introduction of this special navigation language will let us to control the localization and position refinement process; thus, no RF infrastructure is necessary and continuous user supervision is minimized or at least is automatically triggered by the scenario itself.

As future work, we have to put the pieces together and to implement the localization service on the server side. Currently the data gathering server, the ViSe querying on the mobile side and an off-line version of the movement pattern detection algorithms are available. The camera based position refinement is also under implementation and evaluation.

Acknowledgment

The work of Tamás Helfenbein was founded by the National Innovation Office (NIH) via the project BelAmI_H.

Attila Török was supported from project WayFiS. The project WayFiS no AAL-2010-3-014 has received funding from AAL JP, co-funded by the European Commission and National Funding Authorities of country Spain (MINETUR), Switzerland (OPET) and Hungary (NIH).

References

[1] K. Rehrl, S. Bruntsch and H-J. Mentz, ”Assisting Multimodal Travelers: Design and Prototypical Implementation of a Personal Travel Companion,” IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, pp. 31-42, March 2007.

[2] Y. Gu, A. Lo and I. Niemegeers, ”A Survey of Indoor Positioning Systems for Wireless Personal Networks,” IEEE Communications Surveys & Tutorials, vol. 11, no. 1, pp. 13-32, first quarter 2009.

[3] K. Chintalapudi, A.P. Iyer and V.N. Padmanabhan, ”Indoor Localization Without the Pain,” MobiCom’10, September 20-24, 2010, Chicago, Illinois, USA.

[4] Y. Jin, M. Motani, W-S. Soh and J. Zhang, ”SparseTrack: Enhancing Indoor Pedestrian Tracking with Sparse Infrastructure Support,” IEEE Infocom 2010, pp. 668-676, NJ, USA, 2010.

[5] Y. Jin, H-S. Toh, W-S. Soh and W-C. Wong, ”A Robust Dead-Reckoning Pedestrian Tracking System with Low Cost Sensors,” IEEE PerCom, pp. 222 - 230, Seattle, March 21-25, 2011.

[6] Z. Song and et al., ”Dead-Reckoning Assisted WiFi Based Indoor Pedestrian Localization,” ComNet-IoT, January 3, China, 2012.

[7] M. Löchtefeld, S. Gehring, J. Schning and A. Krüger, ”PINwI - Pedestrian Indoor Navigation without Infrastructure,” NordiCHI ’10: Extending Boundaries, pp. 731-734, New York, USA, 2010.

[8] A. Mulloni, H. Seichter and D. Schmalstieg, ”Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions,” MobileHCI 2011, pp. 211-220, Aug 30Sept 2, Stockholm, Sweden, 2011.

[9] A. Mulloni, D. Wagner, I. Barakonyi and D. Schmalstieg, ”Indoor Positioning and Navigation with Camera Phones,” IEEE Pervasive Computing, vol. 8, issue 2, pp. 22-31, 2009.

[10] D. Merico and R. Bisiani, ”Indoor Navigation with Minimal Infrastructure,” Proc. 4th Workshop Positioning, Navigation and Communication (WPNC 07), pp. 141144. 22-22 March, 2007.

[11] C.W. Han, S.J. Kang and N.S. Kim, ”Implementation of HMM-Based Human Activity Recognition Using Single Triaxial Accelerometer,” IEICE Trans. Fundamentals, vol. E93A, no.7, pp. 1379-1383, July 2010.

[12] Q. Pan, C. Arth, E. Rosten, G. Reitmayr and T. Drummond, ”Rapid Scene Reconstruction on Mobile Phones from Panoramic Images,” IEEE ISMAR ’11, pp. 55-64, Washington, DC, USA, 2011.