A Semantic-Enhanced Augmented Reality Tool for OpenStreetMap POI Discovery

Michele Ruta*, Floriano Scioscia, Danilo De Filippis, Saverio Ieva, Mario Binetti, Eugenio Di Sciascio

D.E.I., Politecnico di Bari, via E. Orabona 4, I-70125, Bari, Italy
D.I.C.A.T.E.Ch., Politecnico di Bari, via E. Orabona 4, I-70125, Bari, Italy

Abstract

The Augmented Reality (AR) is often said to have the potential for a revolution in the way we discover Points Of Interest (POIs) and experience our cities. Nevertheless, to date the AR promise has only partially become true, because the information content supporting location-based resource discovery is usually shallow. Semantic-based technologies allow expressing rich, accurate and meaningful descriptions of POIs, so helping in improving the quality of discovery.

Building upon a general framework for the semantic annotation of nodes in the crowd-sourced OpenStreetMap (OSM) cartography, a novel discovery tool in AR is proposed for mobile devices. Based on the user’s personal profile, it shows markers for POIs in the field of sight upon the real-time device camera view. The tool performs automatically a semantic matchmaking between the user profile and the resource descriptions extracted from OSM. Both are expressed according to a common reference ontology. The tool displays the results of matchmaking without user effort, by color-coding the markers. The user can select a marker to see the complete annotated description of the POI as well as matching, missing and conflicting elements with respect to her profile.

A fully functional tool prototype was developed for Android mobile devices. Its context-aware user interface makes advanced discovery practical and seamless. A case study was conducted in the city of Trani of the Apulia region in Italy to assess the effectiveness of the proposal.

* Corresponding author. Tel.: +390805963316; fax: +390805963410.
E-mail address: m.ruta@poliba.it

Keywords: Augmented Reality, OpenStreetMap, Semantic-based matchmaking, Point Of Interest discovery
1. Introduction

Mobile Augmented Reality (MAR) is an active area in mobile computing, enabled by the availability of smart portable devices equipped with Internet connectivity, camera and location sensors (GPS for position, magnetometer for heading, accelerometer and gyroscope for inertial movement). AR enhances the user’s perception of and interaction with the real world, by making it a part of the user interface, so that accessing and understanding location-based information becomes easier. This has led to a widespread adoption in several domains like medical visualization, entertainment, advertising, maintenance and repair, annotation and robot path planning.

Geospatial MAR has been often called a potential revolution in the way we tour our localities, discover Points Of Interest (POIs) and experience our surroundings. Nevertheless, to date the promise has been fulfilled only partially, because the information content supporting location-based resource discovery is usually shallow. POIs are typically characterized only by a name and a category. Such rigid approach only supports exact matches: search cannot exploit more detailed POI attributes to extract and filter resources according to a relevance measure. This often leads to either too many results for broad queries or too few for specific ones. In either case, the MAR user experience suffers.

Semantic-based technologies can allow more articulated and meaningful descriptions of locations and POIs. The use of metadata (annotations) endowed with formal machine-understandable meaning can enable more advanced location-based resource discovery through proper inferences. Building upon previous work (Scioscia, Binetti, Ruta, Ieva, & Di Sciascio, 2013) which leveraged Semantic Web technologies and crowd-sourced OpenStreetMap (OSM, http://www.openstreetmap.org/) cartography to tag POIs with semantic annotations, here a MAR framework and system for semantic-enhanced POI discovery and exploration are proposed. Cartographic data is extracted from semantic-enhanced OSM for the user local area. Starting from an annotated user profile, the system executes matchmaking with semantically annotated POIs in a reference range with respect to user’s position. Outcomes are displayed as color-coded markers on the display used as device camera viewfinder, corresponding to the real direction and distance of each POI from the user. By touching a marker, the user can see a logic-based explanation for the outcome, in terms of missing and/or conflicting characteristics between her profile and the POI. The proposed system can be considered as a general-purpose AR discovery facilitator, because several resource domains can be explored by simply selecting the proper reference ontology. Furthermore, complexity of the underlying logic formalisms and inference services is completely hidden from the user through a friendly GUI. Therefore the usage experience is intuitive and pleasant even for non-expert users.

The remaining of the paper is structured as follows. In next Section, relevant related work is briefly surveyed. The overall framework for semantic-based AR is explained in detail in Section 3. In order to clarify benefits of the proposal, a case study in the city of Trani (Italy) is described in Section 4. Finally, conclusions and future work are reported in Section 5.

2. Related work

(Carmigniani, et al., 2011) published one of the most recent and comprehensive reviews of AR technologies, systems and applications. They discussed seminal works as well as recent developments and perspectives. Besides technical effectiveness, adaptability to contexts and use cases, important open issues concern privacy and social acceptance.

Mobile augmented reality becomes increasingly feasible on mass-market inexpensive hardware, such as smartphones and tablets. For a widespread adoption, a broadly adopted standard is needed for authoring and distributing content, with similar flexibility and interactivity to current Web authoring technologies (Hill, MacIntyre, Gandy, Davidson, & Rouzati, 2010). (Wikitude GmbH, 2014) and (Layar, Inc., 2014) are among the first commercial MAR browsers, which display third-party digital content created via proprietary APIs. (Schmalstieg, et al., 2007) proposed Studiersube, a full MAR modeling pipeline composed of four tiers: model acquisition, storage, delivery to applications, use. The core data model devised for the storage tier is the unifying element of the architecture: an XML application called Building Augmentation Markup Language (BAUML) covers both indoor and outdoor 3D environments, including topological information to derive navigation hints. (Hill, MacIntyre, Gandy, Davidson, & Rouzati, 2010) proposed a MAR application architecture combining KML (Keyhole Markup
Language, the XML application used by Google Earth), HTML5 and JavaScript. Wikitude GmbH was the original proponent of the Augmented Reality Markup Language – ARML (Lechner, 2014), a candidate standard of OGC (Open Geospatial Consortium). It provides: an extensible object model to describe augmented reality applications, scenes and resources; an XML Schema Definition to represent them; ECMAScript bindings for applications to manipulate them. The goal of standardization is to promote the eventual interoperability among asset and solution providers in this field.

Tracking technologies used in MAR include: optical sensors and computer vision (marker-based or markerless), location and orientation sensors (GPS, accelerometers, magnetometers, ultrasound), radio beacons (RFID, Bluetooth, WiFi, Ultra-WideBand) (Carmigniani, et al., 2011). The tool proposed here does not use computer vision, but exploits only location sensors. They are adequate for scenarios where the AR system does not need to mediate manipulation of real-world objects. Furthermore, real-time tracking based on computer vision is still challenging for smartphones, despite the ongoing hardware and software improvements (Morrison, et al., 2009) (Wagner, Reitmayr, Mulloni, Drummond, & Schmalstieg, 2010).

Latest MAR developments focused on innovative applications. Developing successful services around MAR is still in its infancy, which is partially resulting from the lack of insight into potential users’ expectations and acceptance. Early surveys showed that the true values of AR services are in making contextually relevant information easily available and enabling cooperation (Morrison, et al., 2009) (Olsson, Kärkkäinen, Lagerstam, & Ventá-Olkkonen, 2012). (Kovachev, Niculaescu, & Klamma, 2013) proposed a MAR solution for collaborative location-based multimedia editing. They used XMPP (Extensible Messaging and Presence Protocol) for user coordination and MPEG-7 for XML-based multimedia annotations. Similarly, (Lin, Chen, Li, Wu, & Chen, 2013) presented a MAR application for job hunting, supported by a cloud database and integrated with social networks. In both proposals, however, semantics of annotations is quite shallow and does not allow a content-based POI discovery; on the other hand, the collaborative and social features are interesting and are already a part of the proposed MAR framework by leveraging the OpenStreetMap crowd-sourcing.

In the above proposals, POI discovery was not supported by advanced semantic characterization and matchmaking. The only relevant MAR proposal based on Semantic Web technologies is from (Van Aart, Wielinga, & Van Hage, 2010). They presented a MAR client for iPhone which retrieves an RDF (Klyne & Carroll, 2004) dataset relevant to locations and objects in the direction of the user. The proposed framework integrated Linked Data (Bizer, Heath, & Berners-Lee, 2009) sources of cultural heritage collections. The applicability of the approach was limited by the availability of pre-existing RDF datasets, since the problem of creating and maintaining them was not considered. Our previous research efforts for mobile semantic POI discovery were affected by similar issues, which can be overcome by exploiting crowd-sourcing with open map data.

3. Framework architecture

Fig. 1 depicts the overall architecture of the proposed framework. It consists of the following components:

- The OpenStreetMap server working as cartography provider. OSM map entities are semantically enriched in a way that best fits location-based resource discovery, as explained in Section 3.1. The standard OWL 2 Web Ontology Language (W3C OWL Working Group, 2012) languages are exploited to create and share POI annotations, based on ontologies providing the conceptual vocabulary to express them and enabling automated inferences.

- A general method and an editor (Scioscia, Binetti, Ruta, Ieva, & Di Sciascio, 2013) for annotating maps, so allowing a collaborative crowd-sourced enrichment of basic OpenStreetMap cartography.

- A mobile augmented reality client providing the following features: (i) discovery of most relevant POIs w.r.t. user’s semantically annotated profile, via a logic-based matchmaking; (ii) visualization of POI annotations and examination of discovery results, through a fully visual user interface.

The present work focuses on the mobile client. Its main components are summarized as follows:

- AR GUI. It is able to display POI markers on top of the smartphone screen used as camera viewfinder, exploiting embedded GPS, compass, accelerometer and gyroscope sensors.
• **Semantic Matchmaker.** The lightweight matchmaker in (Ruta M., Scioscia, Di Sciascio, Gramegna, & Loseto, 2012) was integrated to compute the semantic matchmaking by means of non-standard inference services.

• **OSM data parser.** User profile is processed by the navigation module, which is the core of the system. It extracts information from installed OSM map files.

• **Map data files.** Map data are locally cached in one or more files, which encapsulate both geographical data and semantic annotations of resources. The file structure is described in detail in the next section.

![Fig. 1 Architecture of the proposed framework](image)

**3.1. Semantic enrichment of OSM maps**

In (Scioscia, Binetti, Ruta, Ieva, & Di Sciascio, 2013) a prototypical software tool was designed and implemented for editing semantic map annotations. Particularly, it has been developed as JOSM (Java OpenStreetMap editor) plugin. In order to allow users to store semantic annotations in a POI description retaining compatibility with the current OSM storage structure, some novel tags have been introduced complying with the basic **key-value pair** pattern of OSM element tags:

```
<tag k="semantic:n:key" v="value" />
```

The semantic prefix is used to distinguish semantic annotations from other tags. The index \( n \) identifies different annotations – possibly referring to different ontologies – associated to the same map element. Key name suffix and value format differ for each proposed tag type, as in what follows:
• `<tag k="semantic:n:ontology" v="URI" />` denotes the ontology the semantic node annotation refers to. The tag value is the unique ontology URI (Uniform Resource Identifier), as recommended by W3C (World Wide Web Consortium) specifications, which usually consists of a URL (Uniform Resource Locator) which can be accessed to retrieve the ontology.

• `<tag k="semantic:n:encoding" v="format" />` specifies the compression format used to encode the semantic annotation. Compression techniques are needed in order to cope with the well-known verbosity of XML-based ontological languages such as RDF and OWL.

• `<tag k="semantic:n:counter" v="data" />` tags contain the Base64 string representation of the compressed semantic annotation. If its length is within 255 characters, a single tag is used, else it is split in 255-character segments and each one is stored in a tag. The counter suffix is assigned as a segment index, starting from 1.

3.2. Semantic-based POI discovery in AR

In order to allow users to exploit enriched maps, a mobile augmented reality explorer software system was developed, which introduces novel advanced functionalities for semantic-based resource discovery in mobility. The client was developed using Android SDK Tools, Revision 23, corresponding to Android Platform version 4.2.2 (API level 17).

The process of semantic matchmaking is sketched in Fig. 2. Standard reasoning services for matchmaking include Subsumption and Satisfiability. Given a request $R$ and an available resource $S$, Subsumption checks whether all features in $R$ are included in $S$: its outcome is either “full match” or not. Satisfiability verifies whether any constraint in $R$ contradicts some specification in $S$, hence it divides resources in “compatible” (a.k.a. potential match) and “incompatible” (a.k.a. partial match) ones w.r.t. a request. This approach usually gives poor results, because full matches seldom occur and incompatibility is frequent when matching articulate descriptions. Using standard inferences one cannot understand what constraints caused incompatibility (or missed full match), nor how much they are truly important for the user. In order to give a finer ranking of potential and partial matches, as well as an explanation of outcomes, Concept Abduction and Concept Contraction non-standard inference services were adapted from their original e-commerce scenarios (Colucci, et al., 2007) to POI retrieval. If compatibility is not satisfied, Contraction detects what part $G$ (for Give up) of $R$ is conflicting with $S$ and what part $K$ (for Keep) is not. If one retracts $G$ from $R$, $K$ is obtained, which represents a contracted version of the original request, such that it is compatible with $S$. Therefore Contraction is an extension and an explanation of (un)Satisfiability. On the other hand, if $R$ and $S$ are compatible, but $S$ does not fully satisfy $R$, Abduction identifies what is missing in $S$ in order to reach a full match. In other words, Abduction provides an explanation for (missed) Subsumption, returning what additional

![Fig. 2 Semantic matchmaking steps](image-url)
feature set $H$ (for *Hypothesis*) should be hypothesized in $S$. Furthermore, *penalty functions* can be associated to $G$ and $H$, in order to compute a semantic distance score of each available resource w.r.t. a given request (Colucci, et al., 2007). In Abduction and Contraction, penalty grows accordingly to the number (and type) of concepts in $H$ and $G$, respectively.

A score is finally given to each POI, expressing the result of the matchmaking between the user profile and the POI. The overall resource score is computed using the utility function:

$$f(S, POI) = 100 \left[ 1 - \frac{s \text{match}(R, POI)}{s \text{match}(R, T)} \left( 1 + \frac{\text{distance}(User \_GPS, POI \_GPS)}{\text{max distance}} \right) \right]$$

where $s \text{match}(R, POI)$ is the semantic distance between profile $R$ and $POI$; this value is normalized dividing by $s \text{match}(R, T)$, which is the distance between $R$ and the universal concept (a.k.a. *Top or Thing*) and it depends only on axioms in the ontology. Geographical distance (normalized by user-specified maximum range) is combined as weighting factor. The purposes of the utility function are to weight the result of semantic matching according to distance and to convert the score to a more user-friendly scale. Of course nearer resources are preferred, but in case of a full match $s \text{match}(R, POI) = 0$ hence $f(R, POI) = 100$ regardless of distance.

In the proposed AR POI discovery framework, the user profile plays the role of request $R$. The profile is either composed by browsing visually the ontology modeling the reference domain (Scioscia, Binetti, Ruta, Ieva, & Di Sciascio, 2013), or imported from other applications and services. Available resources $S$ are the annotated OSM POIs in the user’s area, referring to the same ontology as the user profile. They are extracted from OpenStreetMap server and cached in the MAR client. Several resource domains (cultural heritage, shopping, accommodation, etc.) can be explored by simply selecting the proper reference ontology. Hence the proposed system works as a general-purpose location-based service discovery facilitator.

Once the tool starts, the OSM Data Parser module extracts from the OSM map file the list of points of interest, filtering only those endowed with semantic annotation. Next, for each semantic POI, the following steps are performed:

- chaining of the 255 characters blocks;
- Base64 decoding and generation of the compressed file;
- decompression of the annotation.

The system adopted a modified version of the *Android Augmented Reality* framework†. Given POI target coordinates (latitude, longitude and altitude), it collects the azimuth and inclination angle between the device and the target from gyroscope and compass, in order to calculate where the device is pointing and its degree of tilt. Further correction is applied to the data set to compensate for the difference between true north and magnetic north. Using this knowledge, the system decides if and where a POI marker should be displayed within the viewfinder image on the screen.

The semantic description concerning each POI is stored as an attribute of its marker. Matchmaking outcomes are represented graphically in three subsequent levels of detail:

- color-coding the augmented reality marker: red markers for POIs resulting in partial matches with the request, yellow for potential matches and green for full ones;
- a pie chart in the result list reporting the numerical score;
- a list of missing or conflicting logical elements, displayed by means of icons.

In the next section, practical examples will clarify and explain in depth how the system works.

---

† https://code.google.com/p/android-augment-reality-framework/
4. Case study

In order to test and evaluate the devised framework and tool, a case study in cultural heritage tourism sector was carried out in the city of Trani, Italy, using an off-the-shelf smartphone\(^1\). An excerpt is reported here as a toy example, in order to show the benefits of the proposal.

**Semantic-enhanced POI search.** A tourist is visiting Trani for the first time. She is interested in local architectural works, so she runs the MAR tool on her smartphone. The user interface of the tool is displayed in the screenshot in Fig. 3. It displays on a radar several semantic-enriched points of interest within a radius which is adjustable from a slider on the right hand side. Markers for POIs within the field of sight are also shown upon the real-time device camera view. The user can select among them the one she might want to visit.

In detail, at startup the system executes the following steps:

1. The user’s annotated profile consists of a concept expression. It specifies personal information like interests and hobbies. The user has input it in the first-time setup phase or has imported it from another installed application.
2. Check the presence of the matchmaker, which runs as a background service, listening for reasoning service requests.
3. Start keeping track of user position, by means of available location services, e.g. GPS antenna or platform location APIs.

POI discovery can happen in two different ways.

- User explicitly submits her request selecting related features from main menu. Basically she indicates the general resource domain (e.g., Cultural Heritage, Accommodation, Entertainment, Dining, etc.), which corresponds to an ontology. Furthermore, she browses visually the concepts and roles within the ontology to compose the request.
- Alternatively, the navigation tool is able to retrieve POIs best matching user interests in a transparent fashion simply referring to her profile and preferences.

She has already built her personal profile, specifying the venues and architectural features she is interested in. It describes an ancient church in Romanesque style, built in tuff, with longitudinal floor plan. She has also set a maximum acceptable distance from her current location.

As shown in Fig. 1, the embedded reasoning engine automatically applies matchmaking between user profile and nearby POIs. Their markers appear color-coded on the screen, in order to give the user a quick and clear overview regarding most relevant resources.

Semantic similarity scores are displayed in the score panel – A in Fig. 4a – in which the user can obtain the list of resources along with their overall scores as pie charts. In the proposed example, the cathedral of **St. Nicholas** is the POI that better satisfies the user request; in fact, it almost entirely matches her preferences. The churches of **Madonna del Carmine** and **St. Teresa** have good scores because they have no conflicts w.r.t. the request, since differences in artistic styles or floor plan types are not modelled as conflicts in the reference ontology. The **Hohenstaufen Castle** has a lower score because, being a castle, it is in contrast with the request according to concept definitions in the ontology; nevertheless, some requested features are satisfied.

\(^1\) HTC One X with ARM Cortex A9 Quad Core CPU at 1.5 GHz, 1 GB RAM, 16 GB internal memory, and Android version 4.2.2
Result selection and analysis. The user wants to review search outcomes. Selecting a resource, she can see its relevant features, which are presented as icons around a wheel shape, in order to provide a clear and concise description, as shown in the central portion of Fig. 4(a) and Fig. 4(b). Furthermore, the View result panel (B Fig. 4b) lists all missing features computed through Concept Abduction.

The user selects the church of Madonna del Carmine, which is located in close proximity, so she realizes that it is not an ancient age church. It is instead a modern church, in Romanesque style with neoclassical elements. Nevertheless, it could definitely be a good choice, although it deviates from the profile (which appears whenever the user presses the appropriate button on the right, as shown in C in Fig. 4b) regarding the quadrangular floor plan, the epoch and the material. Alternatively, the same left-hand menu can show features computed through Concept Contraction in case of incompatibility: properties that the POI satisfies, or incompatible elements (Fig. 5).

After visiting the church, the user looks in the radar for the presence of other nearby monuments. Turning towards the harbour of Trani, three other markers appear on the screen, as already shown in Fig. 2. Her next choice falls on the Hohenstaufen Castle. By exploring its features (as in Fig. 5), she can see that the incompatibility suggested by the red marker is due to their different nature. Finally, she decides to visit the cathedral of St. Nicholas the Pilgrim, a Romanesque cathedral, which deviates from the profile only by the age of construction, which dates back to the Middle Ages.
Overall, the user can quickly identify what POI resources are most relevant to her needs and desires, by looking at the POI marker color, at the matchmaking result shown in the score panel and -if interested- by exploring POIs features. Simple operations on the device touchscreen allow effortless information acquisition and management.

5. Conclusion and future work

The paper presented a mobile framework and tool for semantic-enhanced POI discovery in AR. It allows users to see an overlay of markers for points of interest on the scene framed by her mobile device camera. Exploiting semantic-enhanced OpenStreetMap cartography and an embedded lightweight matchmaker, it executes semantic matchmaking between the user profile and the annotations of POIs in her surroundings. Explanation of outcomes is provided at different levels of detail through a set of visual cues, making interaction quick and effortless. A case study in the city of Trani was presented to clarify the novel aspects of the proposal.

Future work includes: managing more articulate POI descriptions through an enhancement of the expressiveness of the underlying logical language; speech-to-text support for voice-based queries, in order to escape the limitation of the single user profile; allowing the composition of OSM POI annotations directly on the mobile device.

Acknowledgments

The work was partially supported by Italian PON project Puglia@Service.

References


