Molé: a Scalable, User-Generated WiFi Positioning Engine

Jonathan Ledlie*, Jun-geun Park**, Dorothy Curtis**, André Cavalcante***, Leonardo Camara***, Afonso Costa*** and Robson Vieira***

*Nokia Research Center, Cambridge, USA. Email: jonathan.ledlie@nokia.com

**Massachusetts Institute of Technology, Cambridge, USA. Email: {jgpark,dcurtis}@mit.edu

***Nokia Institute of Technology, Manaus, Brazil. Email: {andre.cavalcante,leonardo.camara,afonso.costa,robson.domingos}@indt.org.br

***Abstract*—We describe the design, implementation, and evaluation of Molé, a mobile organic localization engine. Unlike previous work on crowd-sourced WiFi positioning, Molé uses a hierarchical name space. By not relying on a map and by being more strict than uninterpreted names for places, Molé aims for a more flexible and scalable point in the design space of localization systems. Molé employs several new techniques, including a new statistical positioning algorithm to differentiate between neighboring places, a motion detector to reduce update lag, and a scalable "cloud"-based fingerprint distribution system. Molé's localization algorithm, called Maximum Overlap (MAO), accounts for temporal variations in a place's fingerprint in a principled manner. It also allows for aggregation of fingerprints from many users and is compact enough for on-device storage. We show through end-to-end experiments in two deployments that MAO is significantly more accurate than state-of-the-art Bayesian-based localizers. We also show that non-experts can use Molé to quickly survey a building, enabling room-grained location-based services for themselves and others.**

I. Introduction

The ability for a mobile device to perceive a user's location has many applications, from social networking "check-ins" to location-appropriate content, such as automatically presenting people with a relevant train schedule.

While the global positioning system (GPS) enables devices to sense their location in most outdoor environments, bad weather and "urban canyons" can restrict its operation. In addition, there are many indoor positioning applications where GPS can provide only limited assistance, as it typically provides a position fix only near windows and doors.

To enable room-grain indoor and outdoor positioning in GPS-less environments, researchers have used physically-fixed wireless beacons to associate a unique "fingerprint" with each place or grid point [1]–[4]. While the types of wireless beacons have varied over time, most techniques now use 802.11 WiFi beacons because of their near ubiquity, particularly in urban and suburban environments. Because of the difficulty in translating between distance and received signal strength [5], more compact alternatives to fingerprinting – e.g., triangulating among the beacons – are generally eschewed.

One of the key problems with fingerprinting, however, is learning the fingerprint for each place – however "places" are designated. We call the process where a person links a fingerprint to a place "binding." Several commercial vendors offer positioning services, which include a fingerprint-generation survey [6]. However, these come at a steep price: a large office building can cost $10,000 USD with no maintenance included. Because this is prohibitively expensive for many applications – such as contextualizing a device's behavior based on which room of a house it is in – several research groups have begun to crowd-source fingerprints from end-users [7]–[10]. In the model for these Wikipedia-style approaches, a single locally-knowledgeable user performs the bind for a place and many visitors can then rely on the database of fingerprints.

Molé focuses on a new point in the design space in crowd-sourced, or "organic," positioning systems. Some systems, such as OIL [8], present a map to the user: users bind places by clicking on the map. Others, like Redpin [7], allow the association of any text string with a place's fingerprint. In contrast, Molé arranges the world hierarchically; this imposes a clean, intuitive namespace (country, region, ...), and allows for data prefetching at a building scale if not larger. It also isolates problems in the fingerprint database to small portions of the tree. Molé relies on compact data structures that allow many fingerprints to be stored on the user's device. In turn, this allows the user's device – not a server – to differentiate among potential places with similar fingerprints, improving privacy.

In our experimental results, which are discussed in more detail in the full version of our paper, Molé's positioning algorithm, MAO, achieves a 10% improvement over the state-of-the-art [3]. In a live, controlled experiment, it had 93.1% spot-on accuracy. In a crowd-sourced experiment with four untrained users, it grew to a similar level of accuracy after an hour's worth of participation and experienced no wrong-floor errors.

II. Model of Places

Molé arranges the discrete, human-designated places of the world in a hierarchy. While the hierarchy could be of variable depth, our current implementation contains five levels, as the estimate in Figure 1 illustrates. From coarse to fine, the levels typically refer to country, region, city, area, and unique place (e.g. room). Areas are the unit of fingerprint aggregation, transfer, and, therefore, privacy; the server knows at most what areas you visit. Areas typically refer to street addresses (e.g., "4 Cambridge Center" in Figure 1), although they could refer

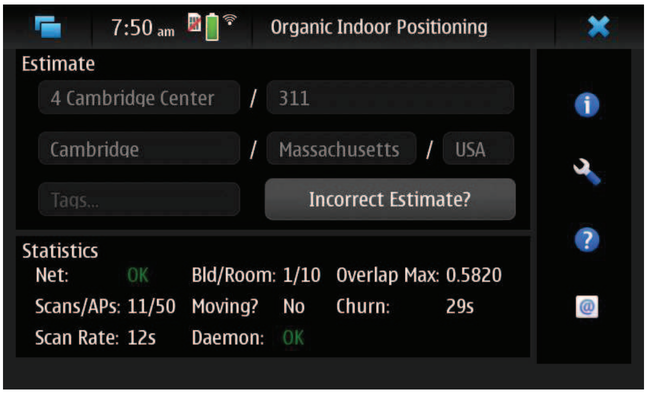

Fig. 1. **Molé's User Interface.** It shows the country, region, city, area, room hierarchy in street address format.

to larger outdoor areas such as parks. The design also allows aggregation at higher levels.

III. Implementation

Molé's implementation is divided into client and server components. The client portion periodically scans WiFi signals and makes an estimate of the current place available to other applications on the same mobile device. Because all position estimates are calculated on the client using a cache of fingerprints, the client's exact position remains private and new estimates can be made in the absence of network connectivity.

Molé's server components are currently hosted on Amazon Web Services[1]. The source code for Molé has been released under an open source license and we invite contributions[2].

A. Client Components

The client itself consists of two parts: a daemon, which runs continuously in the background, and a user interface, which is displayed when the user wants to make a bind, modify the daemon's behavior, or view statistics. Figure 1 shows the user interface. Its statistics include: the number of scans being used to form the estimate; the count of distinct APs that were observed within these scans; the current time between scans (i.e. scan period); the number of areas and individual places within those areas under MAO consideration; whether the user is deemed to be moving; the score of the current estimate ("overlap max"); and churn, the time since the estimate was last changed. The Molé daemon exports the current location estimate to all applications on the device.

Using Motion Detection As Haeberlen et al. showed [3], comparing more user scans against each fingerprint improves spot-on accuracy, with diminishing returns after about eight scans with their data. But frequent scanning reduces battery life, and having a fixed, large number of user scans introduces a lag when the user is moving between places. If a device has an accelerometer, Molé uses it to find a happy medium between battery consumption and update lag. If the device is estimated to be stationary, it slows down the scan rate and

[1]http://mole.research.nokia.com
[2]http://github.com/organic-positioning

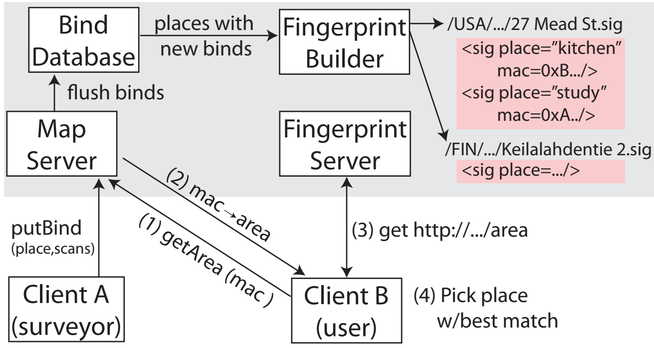

Fig. 2. Interaction between Molé's client and server components. Two paths are shown: (a) a bind coming from a surveyor (client A), being added to the bind database, and being processed into an area's fingerprint file (e.g. Keilalahdentie 2.sig) and (b) a user's device (client B) updating its local cache of fingerprints for the areas that it is potentially in. First it queries to see which areas match a random "loud" MAC with `getArea()`, then it fetches the fingerprint files for those areas. After its cache is up-to-date, it can form a position estimate locally.

other functions. When walking is detected, the current set of user scans is discarded and the scan rate is increased (up to once per 10s in our current implementation). By truncating the user scans (11 in Figure 1), Molé returns a less accurate, but more timely estimate. When the user stops moving, the user scans accumulate and the estimate improves. Because we simply truncate the positioning and bind queues in response to movement, our method is independent of the choice of the particular motion detection algorithm; we use Shafer and Chang's detector [11]. To further reduce battery usage, we run the motion detector every 10 seconds with a duty cycle of 5%; at this rate motion detection has little effect on the overall battery consumption of a typical smartphone.

B. Server Components

Figure 2 shows Molé's four main server components and the key methods clients use to make binds and access fingerprints.

References

[1] P. Bahl and V. N. Padmanabhan, "RADAR: An In-Building RF-Based User Location and Tracking System," in *INFOCOM*, Mar. 2000.
[2] N. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-Support System," in *MOBICOM*, Boston, MA, Aug. 2000.
[3] A. Haeberlen *et al.*, "Practical Robust Localization over Large-Scale 802.11 Wireless Networks," in *MOBICOM*, Philadelphia, PA, Sep. 2004.
[4] J. Krumm and J. Platt, "Minimizing Calibration Effort for an Indoor 802.11 Device Location Measurement System," Microsoft Research, Tech. Rep. MSR-TR-03-82, Nov. 2003.
[5] K. Pahlavan *et al.*, "Wideband Radio Propagation Modeling for Indoor Geolocation Applications," *IEEE Comm. Mag.*, April 1998.
[6] "Ekahau positioning engine," ekahau.com.
[7] P. Bolliger, "RedPin: Adaptive, Zero-Configuration Indoor Localization," in *LoCA*, Oberpfaffenhofen, Germany, 2008.
[8] J. Park *et al.*, "Growing an Organic Indoor Location System," in *MobiSys*, San Francisco, CA, Jun. 2010.
[9] E. S. Bhasker, S. W. Brown *et al.*, "Employing user feedback for fast, accurate, low-maintenance geolocationing," in *PerCom*, Mar. 2004.
[10] A. Barry, B. Fischer, and M. Chang, "A long-duration study of user-trained 802.11 localization," in *MELT*, Orlando, FL, Sep. 2009.
[11] I. Shafer and M. L. Chang, "Movement Detection for Power-Efficient Smartphone WLAN Localization," in *WSWiM*, Miami Beach, FL, 2010.
[12] B. Charrow, "Organic Indoor Location: Infrastructure and Applications," Master's thesis, Massachusetts Institute of Technology, Feb. 2010.