LandLoc: Landmark-based User Location

Venkata N. Padmanabhan
Microsoft Research

January 2001

Technical Report
MSR-TR-2001-23

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

LandLoc: Landmark-based User Location*

Venkata N. Padmanabhan†
Microsoft Research

Abstract

The growing interest in location-aware computing has spurred the development of systems and technologies to locate users, both in indoor environments and the outdoors. These include the global positioning system (GPS), the forthcoming mobile phone based wireless E-911 system, and several indoor location systems. These systems depend on distributed infrastructure support to determine user location, which makes them susceptible to the non-availability or temporary failure of the infrastructure. The cost of the dedicated infrastructure and the potential threat to user privacy in some cases are also drawbacks.

In this paper, we present a novel approach to the user-location problem that requires little or no infrastructural support. We call this approach landmark-based location, or LandLoc for short. The basic idea is that the user identifies landmarks in his or her vicinity and feeds this information into a mobile computer. The system compares the landmark information against a 3D topographical model of the physical world, thereby narrowing in on the user's location. We discuss the technical challenges presented by LandLoc. We do not claim that the LandLoc approach is superior to existing schemes in all respects. Rather we present it as a promising research direction that offers several advantages and may yield a solution complementary to existing systems.

1 Introduction

A fundamental goal of computing is to enable people to interact more effectively with their environment. A user's location is an important aspect of his or her environment. Hence, there has been a growing interest in location-aware computing, which encompasses systems and services that modify their behavior as a function of the user's location [15].

Determining user location is a prerequisite to location-aware computing. This challenging problem has received much attention in recent years. The most developed system for user location in the outdoors is the satellite-based Global Positioning System (GPS) [6]. More recently, there has been an effort on building location capabilities into cellular telephone networks to support emergency relief operations (dubbed *wireless E-911* by the U.S. FCC [7]). For indoor environments, several systems based on wireless networks have been developed. These include infrared (IR) based systems (e.g., Active Badge [8]), radio frequency (RF) based systems (e.g., RADAR [2]), and systems based on a combination of RF and ultrasound (e.g., Active Bats [9]).

A common feature of all of these systems is that they depend on wireless communication between a mobile terminal (such as a laptop or a cell phone) carried by the user and a fixed infrastructure. User location is determined using a variety of measurements on the timing, phase and/or strength of the wireless signal. Such an infrastructure-based approach has several limitations:

1. The infrastructure is deployed solely for the purpose of user location. Examples include satellites in the case of GPS and infrared transmitters in the case of the Active Badge system. In addition, specialized equipment is needed at the user/mobile end. Deploying and operating the dedicated infrastructure may be expensive.

2. Dependence on the infrastructure leaves the user location system vulnerable to failure when the infrastructure is either unavailable or temporarily unreachable. GPS does not work indoors or when there is occlusion due to tall buildings or foliage; the walls and furniture in a building block infrared signals; a mobile phone based system does not work where there is no cellular coverage.

3. In some systems, the infrastruture determines the location of users. So there is the potential for privacy violation[1] Not all infrastructure-based systems suffer from this problem because in some cases (e.g., GPS) the mobile device carried by the user determines its own location by being a passive listener.

4. The infrastructure operator controls every aspect of the system's functioning, leaving little control in the

---

[*] This paper presents speculative ideas for a new approach to mobile user location. These ideas are still being refined and have not yet been validated through experiments or prototypes.

[†] http://www.research.microsoft.com/~padmanab/

[1] It is may be possible to ensure privacy by hiding the *identity* of the user from the infrastructure. However, in some cases (e.g., E-911) the user's identity may be exposed to the infrastructure for reasons such as billing.

hands of local organizations such as the city government or a university campus. For example, the U.S. military, which runs the GPS system, can choose to degrade accuracy of the location estimate available to users globally using a feature called *selective availability.*

In this paper, we present a novel approach to user location determination that avoids the limitations listed above. Our scheme, which we call *landmark-based user location* or *LandLoc* for short, is inspired by the approach taken by people while finding their way in an unfamiliar city. For example, a person trying to find a restaurant might call up a friend and say something like: "I just drove past a Shell gas station on the right and now I see a post office on the left. Am I headed the right way?". In other words, the person tries to find the way by identifying *landmarks* in his or her vicinity and querying the friend who is presumably familiar with the area (and the landmarks).

LandLoc operates in a similar manner (Figure 1). The user identifies landmarks in his or her vicinity and feeds this information into a computer. A landmark could be any distinctive object such as a building, a shop, a park, a bridge, etc. The computer, which takes the place of the friend in the above scenario, first locates these landmarks in a 3-dimensional electronic representation of the physical space, such as a city or a mall. It then estimates the user's location based on the location of the landmarks and its knowledge of the topography of region the user is in.

The key advantage of the LandLoc approach is that it is not dependent on any infrastructure support. An electronic model of the topography can be stored on a mobile computing device carried by the user and the computing needed to determine location could be performed on the same device. This infrastructure-independence offers several advantages:

1. The cost of deploying a distributed infrastructure solely for the purpose of locating users is avoided. There is a cost associated with creating and maintaining a topographical model. However, we believe it is likely that in the future 3D models of cities, buildings, malls, etc. will be created to support applications such as virtual tours, urban planning and disaster management. So LandLoc could just use these models "for free".

2. Since there is no dependence on an infrastructure, the availability of the system is not impacted by failures of the infrastructure.

3. User privacy is maintained since the user only communicates with his or her mobile computer. There is no (wireless) network communication.

4. Local authorities such as the city government or the owner of a building have control over the accuracy of the topographical model available to the user.

The LandLoc approach presents several challenges. These include constructing a topography model efficiently, storing and querying the topography model on the user's mobile device, and providing the user with a convenient interface to input landmark information while also resolving ambiguities in the user's identification of landmarks (for example, the user may identify a "tall building", but there may be several such buildings in a city).

We believe that the promise of the LandLoc approach to determining user location coupled with the challenges mentioned above make it a promising avenue of future research. An effort to building an effective user location system based on LandLoc would draw upon knowledge and research expertise from a diverse set of areas, including mobile computing, computer graphics, computer vision, artificial intelligence, image processing, and speech recognition. We do not claim that the LandLoc approach is superior to existing schemes in all respects. Rather we view it as a promising approach that offers several advantages and may yield a solution that complements existing systems.

The remainder of this paper is organized as follows. In Section 2, we survey several user location systems, pointing out the stengths and shortcomings of each. In Section 3, we outline the proposed LandLoc approach. We discuss various challenges and potential solutions. Finally, we present our conclusions in Section 4.

2 Related Work

We survey several user location systems, targeted both at indoor environments and at the outdoors. We point out the strengths and limitations of these systems.

The *Global Positioning System (GPS)* [6] is by far the most widely deployed user location system. The GPS infrastructure consists of a constellation of 24 satellites and 5 ground support stations spread across the globe. The user is equipped with a special GPS receiver that locks onto radio signals from multiple satellites. A lock on at least 4 satellites is needed for the receiver to be able to estimate its location. The main strengths of GPS are its global reach and its high accuracy (under 10 meters). However, the dependence on the satellite signals means that the system does not work indoors. Tall buildings lining city streets (forming the so-called *urban canyons*) or foliage also make it difficult for the GPS receiver to lock on to the (weak) signals from the requisite number of satellites. Even in the open, it can take several minutes for the first acquision of the GPS satellites' signals by the receiver, which may be a problem in applications where the receiver is only used intermittently.

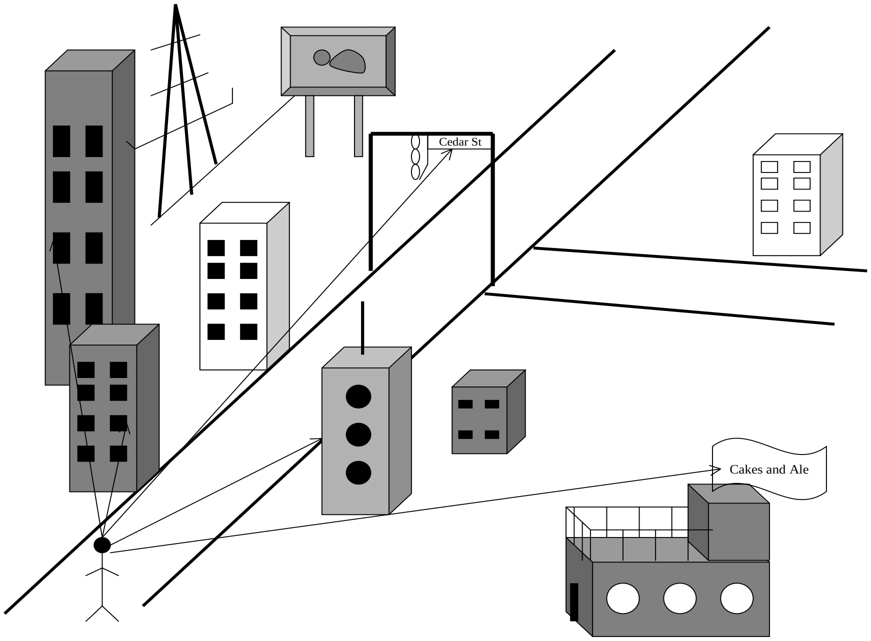

Figure 1: **An urban scenario showing LandLoc in operation. From his present location, the user is able to see two buildings on the left, the one with circular windows on the right, the restaurant on the far right, and the traffic intersection straight ahead (though perhaps not the name of the street). These are the landmarks he would report to the LandLoc system.**

Another approach to user location in the wide area is the *wireless E-911* [7] work undertaken by wireless carriers in the U.S. in response to a directive from the U.S. FCC. This service, which aims at providing a location resolution of 50-300 meters, is scheduled to be operational in late 2001. Several alternative technologies are being considered, including angle of *arrival (AOA)*, *time of arrival (TOA)*, and *time difference of arrival (TDOA)*. By measuring the angle or time delay with respect to multiple base stations, the mobile phone determines its location using trilateration. However, these approaches present several challenges. AOA requires the installation of directional antennas or antenna arrays on the mobile phone (and at the base stations), which is hard to do. The time-based approaches require the addition of special location-determination equipment (called the *location measurement unit (LMU)* in GSM networks), primarily to provide time synchronization. A more fundamental problem is that the mobile needs at least three "hearable" bases stations with "good" geometry (i.e., spread uniformly). Power control techniques, employed to minimize interference in cellular networks, make the detection of neighboring base stations difficult. Finally, these techniques tend to be either purely network-based or network-based with mobile assistance. In either case, the network operator has knowledge of the user's location, which may be problematic from a privacy viewpoint.

Neither GPS nor wireless E-911 is an option for user location indoors. Hence several specialized systems have been developed for indoor environments. The *Active Badge* system [8] uses infrared (IR) beacons emitted by the active badge device to track the user. However, IR requires line of sight between the transmitter and the receiver, so it is blocked by solid objects such as walls and furniture. Direct sunlight, such as through a window, is also a problem. The *RADAR* system [2] uses signal strength measurements from multiple wireless LAN base stations to determine user location. As in E-911 systems, the requirement that at least 3 base stations be within range may not always be possible to satisfy. The *Active Bat* system [9] uses a combination of radio frequency (RF) and ultrasound signals to facilitate the measurement of the distance of the mobile from multiple reference points. The (fast) RF signal is used to synchronize time while the (slower) ultrasound signal is used to measure the distance. Like the Active Badge system, Active Bat requires specialized hardware infrastructure solely for the purpose of user location. Furthermore, both systems "expose" the mobile device by requiring it to transmit, although there is no fundamental reason why this must be so. For instance, the *Cricket* system [12], which is similar to Active Bats in its use of RF and ultrasound (including specialized hardware), eliminates the privacy problem by having the mobile operate purely as a receiver.

We are aware of a couple of systems that are closer in spirit to LandLoc. The first is *Vindigo* [16], a commercial product that provides users with information on points of interest (restaurants, shops, etc.) in the vicinity of their current location. The user downloads a map and database for the city ahead of time. The user can maintain his/her privacy since there is no need for network communication. The user identifies his or her location by specifying the cross-streets at the intersection closest to where he or she is. Therein lies the main difference between the approach Vindigo takes and the one we advocate for LandLoc. The former takes a very restrictive view of what a landmark is. There may be situations where street information is not available (for instance, there may be no street signs, or they may present but in an unfamiliar language or illegible to a person driving down fast). In such situations, the Vindigo approach would ignore other perfectly good landmarks such as buildings, towers, and bridges, whereas LandLoc would make use of such information.

The final system we discuss is *GUIDE* [4], which provides a location-aware tourist guide. The primary user location technique employed by GUIDE is based on a wireless LAN. However, in disconnected mode, the system depends on user assistance to determine location. The user is shown a series of thumbnail pictures of landmarks in the vicinity of his/her last known location. The user identifies the one closest to his/her current location and informs the system. While this approach may work well for a handful of tourist landmarks, it is unlikely to scale to a city with literally thousands of landmarks since the user's task would become onerous.

3 Landmark-based Location of Users

Having surveyed the large body of related work, we now turn to LandLoc, the landmark-based user location approach that we are proposing. Two fundamental goals of LandLoc are: (a) minimal dependence on the infrastructure independence, and (b) complete privacy of the user. In other words, LandLoc aims to let users determine their own location without other entities being involved in any way.

As outlined in Section 1, the operation of LandLoc involves the following basic steps:

1. A 3D topographical model of the region of interest (e.g., a mall or a city) is created.
2. This model is loaded onto the user's mobile computer, in much the same manner as Vindigo does today.
3. To determine his or her location, the user identifies landmarks in the vicinity and feeds this information into his or her mobile computer.
4. The computer estimates the user's location based on its knowledge of the topography of the region and the landmarks identified by the user.

In the following sections we elaborate on these individual steps, pointing out the hard problems to be solved and outlining potential solutions.

3.1 Building a 3D Topographical Model

The primary reason for building a 3D topographical model is to enable LandLoc to accurately estimate the user's location based on the landmarks that he or she is able to *see*. As such we need to model any objects that are either likely to be identified as landmarks or are large enough to cause occlusion.

There has been extensive research in the computer graphics and vision communities on the problem of constructing 3D models of the physical world in an automated fashion. There have been efforts to construct such models for indoor spaces as well as the outdoors. The most promising approach is to use a large collection of photographs to deduce the physical structure of a scene. Prior information, such as a CAD model of a building or a city, can facilitate construction of the model. Examples of research efforts in this space are the Virtual L.A. project [17] and the MIT City Scanning Project [10].

The construction of a 3D model involves several steps, as articulated in [10]:

1. **Capturing images:** Several overlapping photographs of the scene of interest are taken. Photographs are taken from several locations. The position and orientation of the camera is recorded each time a photograph is taken.

2. **Constructing image mosaic:** The individual images are aligned to form a seamless mosaic. Knowledge of the camera position corresponding to each image greatly facilitates this task. A mosaic offers a much larger field of view than any individual image, which results in better efficiency and robustness.

3. **Extracting structure information:** Perhaps the most challenging step of the entire process is reconstrucing structures, such as buildings, from a large number of images. This is an area of active research, with several approaches being investigated. We discuss some of these next. As before, knowledge of camera position is very helpful because it reduces the number of unknowns that the reconstruction algorithm has to deal with.

Several algorithms for 3D reconstruction are dicussed in [10]. Some algorithms are *feature-based*, i.e., they attempt to infer structure based on proximity of features such as corners or edges. Others are *region-based*. They attempt to infer structure by identifying a large number of pixels that appear to lie on the same surface. For example, the facade of a building may be identified in this manner. Yet another approach is to use *voxels* (*volume pixels*, the 3D equivalent of pixels) and multi-scale image processing to extract a volumetric representation of the scene.

A very different reconstruction algorithm presented in [10] uses no feature correspondence. Instead, it uses *edge histogramming* to identify and localize prominent vertical facades under the assumption that such facades exhibit many horizontal edges. The histogramming technique produces a peak whenever many edges reinforce a single facade, yielding the orientations of the dominant facades in the scene. This technique yields facade *geometry* but not texture information. While the latter may be inconsistent across multiple images (each of which may be partially obsured by scene clutter), geometry information correlates across images, so the multiple images reinforce each other to produce a consistent and high-confidence estimate of the shape of a structure such as a building. This is a significant advantage of this technique and makes it particularly relevant to LandLoc, as we discuss next.

LandLoc uses 3D models only to determine the position and orientation of landmarks and to decide on visibility questions (e.g., could the user have seen landmark L from position X?). As such the model only need capture the *geometry* of structures in the scene. For example, it may be sufficient to model a building as a box with four vertical faces, so an approach such as edge histogramming would be a good match. It is not a serious problem if this simplification results in an underestimation of the occlusion caused by the building, but overestimation may be problematic (because it may cause LandLoc to think that landmark L is *not* visible from location X whereas it actually is). Besides geometric information, the model must contain sufficient detail to enable the correlation of landmarks identified by the user with structures in the model. We discuss this issue further in Section 3.3.

One question is how up-to-date the 3D models need to be since the landscape of a city or the layout of a mall can change over time. As we discuss in Section 3.3, LandLoc does not make binary decisions based on the presence or absence of individual landmarks. Rather it considers an ensemble of inputs from the user before making a determination of location. So LandLoc would be robust to occasional errors caused by staleness of the 3D model. Of course, it would be desirable to update the model regularly, both for use with LandLoc and for other applications such as virtual tours. Automated systems, such as [10], would be very helpful.

In summary, it is clear that constructing a 3D model of a physical scene is a challenging problem. However, this is an area of active research and rapid progress in the state-of-the-art. At the present time, such models exist only for limited areas (e.g., the San Francisco bay area), not enough to support LandLoc. However, we expect 3D models of cities, buildings, etc. to become commonplace in the future. The driving applications would be urban planning, virtual tours and walkthroughs, disaster management, etc. LandLoc would then be able to make use of such models "for free".

3.2 Creating a Portable 3D Model

LandLoc requires a 3D topographical model of the region of interest to be stored on the user's mobile computer so that it is remains accessible while the user is "on the road" with no network connection. However, the topographical model could be very large. For example, in the Virtual L.A. project [17], a detailed model of the entire Los Angeles basin (an area of over 10,000 square miles) is being created. The model, when complete, will exceed one terabyte in size. Making a model as large as this portable is clearly a challenge.

We suggest three ways of addressing this challenge:

1. **Compact model:** As mentioned in Section 3.1, LandLoc uses the 3D model primarily to locate landmarks and infer visibility. So a basic model that captures the essential geometric information may be sufficient. Such a model would be quite compact. As a simple example, consider a (large) city area of 100x100 blocks[2] (i.e., 10,000 blocks in all), where each block contains 100 buildings. Suppose a building could be modeled using 4 polygons, one for each face. Then a model of the entire city would contain 4 million polygons, which is certainly a manageable size.

2. **On-demand download:** Another approach would be to partition the model into sections corresponding to geographically contiguous areas. In a large city, there could be sections corresponding to the financial district, the shopping district, etc. Each building or mall could have a separate model for its indoors. The user's mobile computer would hold only the section of the model corresponding to the area he or she is currently in. When the user moves out of one area into the next, new sections are downloaded on demand. Clearly, the network communication needed to download new sections of the model might expose the user to the network infrastructure, potentially resulting in privacy violation. This problem can be avoided by using an approach based on *broadcast disks* [1]. At strategic locations, such as airports and the entrances of shopping malls, the model information could be broadcast on an continuous basis. The user's computer, equipped with a short-range radio such as BlueTooth [3], could receive and store the model information by *silently* listening to the transmission.

3. **Larger mobile storage:** Rapid strides in storage technology may make the storage challenge moot

---

[2] A *block* denotes the area enclosed by adjacent sets of parallel streets that form a grid

within a few years. Historically, disk storage for a fixed cost has been more than doubling each year, with even more rapid progress in recent years. At the present time (early 2001), it is quite common for laptop computers to have disk storage in excess of 10 GB. So it is quite likely that the amount of storage available on mobile computers would be of the order of a terabyte within a short 5-7 years. As mentioned earlier, a significant motivation for generating 3D models of physical spaces is to enable applications such as virtual tours and walkthroughs. For example, given a choice of multiple malls, the user may do a quick virtual walkthrough of each and then decide which one to visit. Unlike LandLoc, a virtual walkthrough application would need to store a detailed model, so the larger storage would come in handy.

3.3 Identifying Landmarks

Next we turn to the problem of landmark identification, both by the user who is trying to determine his/her location and by the LandLoc system, which processes the user's input. The main challenge is that the landmark identification process needs to balance two conflicting requirements: (a) it needs to minimize the burden on the user, but (b) it also needs to remain accurate enough for the LandLoc system to be able to match the user identified landmarks with objects in the model. At the one extreme, the user may identify a landmark as simply "a building". While this minimzes effort on the user's part, the inherent ambiguity makes it difficult for LandLoc to make use of this information. At the other extreme, the user could read out the street name and the number of the building. Such information would be very useful to LandLoc, but obtaining it might be very burdensome for the user.

We need to strike a middle ground. Here are some guidelines that may facilitate this:

1. **Specific characteristics:** It is helpful if the user identifies (some) specific characteristics of a landmark, e.g., its color, its approximate height (a building could be classified as low, multi-storeyed, or a skyskraper), peculiar characteristics (round windows, antenna on top), etc. The user may be able to tell whether he or she is viewing the front of a building or one of its sides. An indication of the direction in which the landmark is present would also help (e.g., if the user says "landmark X is on my left and Y is on my right", LandLoc could narrow down the user's location to the region sandwiched between the two landmarks).

2. **Closer is better:** In general, it is more useful if the user were to identify landmarks that are close to his or her current location than ones that are farther away. Distant landmarks would inherently tend to be visible from points in a large area (e.g., a tall and prominent tower may be visible from much of the locations in a city). As such, knowing that a user can see such a landmark may not help narrow down his or her location very much.

3. **Multiple landmarks:** Given the inherent ambiguity in the identification of landmarks by the user, it would help greatly if the user were to identify multiple landmarks (albeit ambiguously), possibly in different directions. LandLoc could narrow down the user's location when it is presented with an ensemble of (potentially ambiguous) landmark information. For example, the landmarks "large water fountain" and "red building" may, individually, not help much because there could be several large water fountains and red buildings in the city. However, if the user were to specify both, that would help narrow down his or her location because the instances of a water fountain and a red building in the same vicinity may be far less common.

4. **Partial information:** As is clear from our discussion thus far, LandLoc does not require users to provide "complete" information on landmarks (which would be impractical in any case). LandLoc works with bits and pieces of information that the user provides. In other words, it operates on an ensemble of *partial* but *specific* information. For example, the user might specify the color of one of the buildings within sight and also specify that another building in sight has a tall antenna on top. Neither building is completely specified, but taken together these pieces of partial information may be very useful. As another example, suppose there is a building, with the street address "1920 Cedar Street", that has the number 1920 displayed prominently in front. A user driving by may easily be able to notice (and report) the number 1920 but may not know which street he or she is on. Still the partial information, i.e., the number 1920, may be useful to LandLoc in conjunction with information on other landmarks.

5. **Multimedia information:** Part of the information on landmarks can be gathered automatically without explicit user involvement. For example, the user's mobile computer may be equipped with cameras that can capture pictures of the surroundings. The pictures can supplement speech and other input from the user. Such an approach has some similarities to vision-based tracking systems (e.g., [11]), except that here the camera input is only one among many possible inputs and the infrastructure does *not* track the user.

Another important issue is the interface provided by LandLoc to a user who wants to identify landmarks in his or her vicinity. Given our focus on mobile users, we believe that the most appropriate interface would be one based on speech, both for input and output. The rapid improvement in speech recognition technology in recent years is a key enabler. Two aspects of our setting simplify the speech recognition task: (a) the user is likely to use a limited vocabulary when identifying landmarks, and (b) the user is unlikely to use continuous speech, so the challenge of recognizing continuous, conversational speech is avoided.

Identifying landmarks can be an interactive process involving the user and the LandLoc system. Once the user has specified information on landmarks in the vicinity, the system may ask for clarifications. For example, the user may specify a landmark as "a tall building on the right". The system could turn around and ask "what color?" or "do you see a TV tower straight ahead?". Whether and what questions the system asks depend on the impact the additional information will have on LandLoc's ability to narrow down the user's location. For example, based on the information provided by the user, LandLoc may have determined that the user is at one of two locations. To decide between the two locations, LandLoc may decide to ask the user whether a specific building has a mailbox in front of it. Asking for the color of the building may be of no use in this example, perhaps because both buildings are similarly colored.

On a final note, the manner in which users identify landmarks has implications for the kinds of information that the 3D topographical model must contain. In Section 3.1, we stated that the model should capture geometry information so that it can resolve visibility issues. In addition to geometry, model should also include information such as the dominant color(s) of a structure, the shape and/or size of windows, and other peculiar characteristics of the structure. Some of this information (e.g., the presence of an antenna atop a building) may be detected automatically using techniques such as those discussed in Section 3.1. Other information (e.g., the presence of a restaurant on the ground floor of a building) may need to be added in through manual annotation. An experimental field study or user survey would be needed to determine the characteristics of structures that users are likely to identify, and which therefore are important to model. While larger than the simple geometric model discussed earlier, an annotated model is still orders of magnitude more compact than virtual reality models such as the Virtual L.A. model [17].

3.4 Determining User Location

We finally turn to the problem of determining a user's location given both a 3D topographic model of the region (with annotations, as discussed above) and the user's input on the landmarks in his or her vicinity. There are two related problems to be solved: (a) matching the landmarks specified by the user with objects and structures in the model, and (b) given the identity of the objects in the model, determining the location of the user based on visibility information. While we could treat these as logically separate problems, in practice it would be more effective to consider them together owing to the ambiguity inherent in the identification of landmarks by the user.

Conceptually, the procedure is simple. Based on user input, the system draws up a list of "candidate" landmarks that the user might be refering to. For each candidate landmark, the system determines the region (which we call the *visibility region*) from which that landmark would be visible. It then considers various combinations of landmarks and for each combination computes intersection of their visibility regions. Most of the intersections would be empty, indicating that the corresponding combination of landmarks is not valid (for example, if the user reports seeing *both* a post office and a park, it is quite impossible for the post office to be located at the north end of town and the park to be located at the south end, several kilometers apart). The system can disambiguate between the (presumably few) non-empty intersections by asking the user for clarifications (as discussed in Section 3.3).

While conceptually simple, the procedure outline above is likely to be very inefficient because a lot of the computation (both to determine visibility as well as to do the intersection) may be unnecessary and wasteful. For example, it may be possible to eliminate several combinations of the candidate landmarks (such as the post office and park above) using simple heuristics, such as one based on distance (in the above example, it is extremely improbable that the user to see both the post office and the park given that they are both low "structures" and are several kilometers apart, so the combination can be eliminated from consideratio right away). We suggest several possibilities for making the process efficient:

1. **Consider specific information first:** The user may have specified two of the landmarks as "post office" and "tall red building". There may be a lot fewer tall red buildings in town than post offices. Rather than treat all landmarks as "equal", it would be be would be more efficient to first identify the "rare" landmarks, i.e., the red buildings, and then only consider combinations of these with other candidate landmarks in the vicinity. Similarly, it would be best to consider landmarks with a small (i.e., most specific) visibility region first. These procedures are akin to *query optimization* in database systems, where the most restrictive predicate is evaluated first in order to cut down the search space to greatest extent.

2. **Exploit distance and direction information:** In

some cases the user may provide an indication of the distance to a landmark or the (relative) directions of multiple landmarks. For instance, the user might report "a school building on the left" and "a tower a few blocks down on the right". The distance and direction information may help prune the search space of candidate landmarks significantly. It may also help prune the intersection space. For instance, if the user mentions seeing "the front of the Sheraton hotel", that cuts the the search space in half. In general, it is more helpful if the user were to report seeing "surface X" (e.g., the front of the building) rather than just "point X" (e.g., the building itself).

While it is conceptually simple to determine the visibility regions corresponding to two landmarks and then compute their intersection, this is likely to be very expensive computationally due to the irregular shape of the visibility regions in the real world. A *hierarchical* approach [14] would be far more efficient. Here is an outline of this approach. The entire region (e.g., a city or a mall) is progressively sub-divided into smaller sub-regions. For example, at the highest level of the hierarchy is the entire city. The next level of the hierarchy could contain 4 quadrants — northeast, northwest, southeast, and southwest. Each quadrant could be further subdivided. Given landmark X, we determine which quadrant(s) it *may be* visible from (X may be visible from a quadrant of there is at least one point in the quadrant from where X can be seen). It may turn out, for instance, that X is only visible from one of the 4 quadrants. This would cut down the search space for visibility and intersection computations by a factor of 4. After working through a few levels down the hierarchy, the search space may be small enough for a direct computation of the visibility region and intersection.

Determining whether a point (such as landmark X) is visible from any point in a region (such as a quadrant) is a challenging problem, but one that has been the focus of much recent research in the computer graphics and vision communities [5, 13]. There this is termed as the *volume visibility* problem, i.e., the detection of occluded portions of space as seen from a given region. While the details of these techniques is beyond the scope of this paper, the basic idea is simple: rather than consider each occluder (such as a building) individually, multiple occluders are combined together (a process termed *occluder fusion*). A simple example is the leaves of a tree. While each leaf ooccludes only a small portion of the view, the ensemble of leaves may cause significant occlusion, so they are *fused* together. One point to note is that this technique may not always yield a definitive answer on whether landmark X may be visible from a certain region. That is okay because it only means that the region in question should not be pruned from the search space. For efficiency reasons, volume visibility for prominent landmarks could be precomputed offline and the results cached.

4 Conclusions

In this paper, we have presented a novel approach to the user location problem, which we call landmark-based location, or LandLoc. LandLoc combines a 3D topographical model of the region of interest with user-identified landmark information to narrow in on the location of the user. LandLoc applies equally to indoor and outdoor environments (e.g., a city, a mall, or an office building). The key benefits of LandLoc compared to existing approaches are: (a) it does not depend on a distributed infrastructure (such as satellites, terrestrial wireless transmitters, etc.), hence it is immune to the lack of availability and/or failure of the infrastructure, and (b) it allows the user to maintain total privacy with regard to his or her location, by not requiring communication with the outside world.

LandLoc presents several interesting and challenging problems in a wide variety of areas, including mobile computing, computer graphics, speech recognition, and artificial intelligence. In the paper we have pointed out many of these challenges and outlined possible solutions. We are presently in the process of refining these ideas. We believe that LandLoc points to a promising and interesting direction of future research in the area of location-aware computing.

Thus far in the paper, we have always presented LandLoc as a standalone system. However, it is certainly possible to use it in conjunction with other systems. Here are a few examples:

1. LandLoc could be used to complement GPS in regions where GPS does not work (e.g., indoors, urban canyons). Location information from the recent past (obtained from GPS before the user enters an urban canyon) could be used to prune the search space for LandLoc.

2. A user equipped with a thin client (such as a cell phone) would not be able to carry the LandLoc system with him or her. Such a user could still use LandLoc by speaking to a server in the wired network. Clearly, the user will not have the privacy benefit of the full-fledged LandLoc system. However, he/she would still benefit from the coverage and the (relative) infrastructure independence of LandLoc.

Acknowledgements

Michael Cohen, John Snyder, and Rick Szeliski provided useful information and feedback on 3D modeling and visibility issues. John Snyder suggested the hierarchical approach to determing visibility regions. We would like to thank them all.

References

[1] S. Acharya, R. Alonso, M. J. Franklin, and S. Zdonik, Broadcast Disks: Data Management for Asymmetric Communications Environments, *ACM SIGMOD*, June 1995.

[2] P. Bahl and V. N. Padmanabhan. RADAR: An In-Building RF-Based User Location and Tracking System. *IEEE INFOCOM 2000*, March 2000.

[3] BlueTooth SIG, *http://www.bluetooth.org/*.

[4] K. Cheverst, N. Davies, K. Mitchell, and A. Friday, Experiences of Developing and Deploying a Context-Aware Tourist Guide: The GUIDE Project, *ACM Mobicom*, August 2000.

[5] F. Durand, G. Drettakis, J. Thollot, and C. Puech, Conservative Visibility Preprocessing using Extended Projections, *ACM SIGGRAPH*, July 2000.

[6] P. Enge and P. Misra, Special Issue on GPS: The Global Positioning System, *Proc. of the IEEE*, January 1999.

[7] Enhanced 911, U.S. FCC, *http://www.fcc.gov/e911/*

[8] Andy Harter and Andy Hopper, A Distributed Location System for the Active Office, *IEEE Network* Vol.8 No.1, Jan 1994.

[9] A. Harter, A. Hopper, P. Steggles, A. Ward, and P. Webster, The Anatomy of a Context-Aware Application, *Proc. ACM/IEEE Mobicom*, August 1999.

[10] MIT City Scanning Project, *http://city.lcs.mit.edu/city.html*

[11] MSR EasyLiving Project, *http://research.microsoft.com/easyliving/*

[12] N. Priyantha, A. Chakraborty, and H. Balakrishnan, The Cricket Location-Support System, *ACM Mobicom*, August 2000.

[13] G. Schaufler, J. Dorsey, X. Decoret, and F. X. Sillion, Conservative Volumetric Visibility with Occluder Fusion, *ACM SIGGRAPH*, July 2000.

[14] J. Snyder, Personal Communication, January 2001.

[15] M. Weiser, Some Computer Science Issues in Ubiquitous Computing, *CACM*, 36(7), 1993

[16] Vindigo, *http://www.vindigo.com/*

[17] Virtual L.A. Project, *http://www.gsaup.ucla.edu/bill/LA.html.*