Ambient Beacon Localization: Using Sensed Characteristics of the Physical World to Localize Mobile Sensors

Nicholas D. Lane, Hong Lu, Andrew T. Campbell
Computer Science, Dartmouth College
{niclane,hong,campbell}@cs.dartmouth.edu

ABSTRACT

There is a growing need to support localization in low-power mobile sensor networks, both indoors and outdoors, when mobile sensor nodes (e.g., mote class) are incapable of independently estimating their location (e.g., when GPS is inappropriate or too costly), or are unable to leverage localization schemes designed for static sensor networks. To address this challenge, we propose *ambient beacon localization (ABL)*, an unconventional approach that allows mobile sensors to localize by exploiting their ambient physical environment. Ambient beacon localization combines machine learning and free range beacon-based techniques to bind distinct characteristics of the physical world that appear in sensor data of known locations, which we call ambient beacon points (ABPs). Supervised learning algorithms are used to allow mobile sensors to recognize ABPs, i.e., those physical locations that are sufficiently distinguishable in terms of sensed data from the rest of the sensor field. Ambient beacon localization leverages the very same sensed data that nodes are already collecting on behalf of applications. When a mobile sensor finds itself at an ambient beacon point it starts to beacon that location so that other nodes in range of an ambient beacon can localize themselves, for example, by applying existing beacon based localization schemes. In this paper, we present the design of ambient beacon localization and its initial evaluation in a building-sized testbed. Our work is at an early stage but our experimental testbed and simulation results demonstrate that this unusual approach to localization shows promise.

Categories and Subject Descriptors

C.2.4 [**Computer Communication Networks**]: Distributed Systems; C.3 [**Special Purpose and Application Based Systems**]: Real-time and embedded systems

General Terms

Design, Algorithms, Experimentation

EmNets'07, June 25-26, 2007, Cork, Ireland

Keywords

Wireless Sensor Networks, Localization, Mobile Sensor Networks, Machine Learning

1. INTRODUCTION

Localizing sensor data is a fundamental requirement for many sensor network applications, providing the necessary context to interpret the collected sensor data. Much of the work on localization in sensor networks found in the literature [5] [6] [7] assumes static nodes that use multihop communications. The performance of these schemes are typically sensitive to dynamic topology changes and disconnection issues. Furthermore, a new class of sensor networks based on mobile nodes [2] [3] [4] [1] results in frequent topology changes and disconnection, making the use of existing sensor networks localization schemes impractical. To study the sensitivity of existing localization schemes to mobility, we evaluate three related beacon-based localization algorithms: the Monte Carlo Localization (MCL) [5], which is designed for mobile sensor networks, and the Amorphous [6] and Centroid [7] schemes, which are representative of two major classes of algorithms (i.e., multihop distance-vector based and single hop RF connectivity based, respectively) used by a number of existing sensor network localization systems. We use simulation to study the performance of these schemes where node mobility is based on empirical ZebraNet [3] traces, one of the few examples of a deployed mobile sensor network with publicly available data. Unlike the actual ZebraNet deployment, in the simulation only a fraction of the nodes (i.e., beacons) have an external means to self-localize. Figure 1 shows the proportion of failed localization requests from nodes during each simulation run plotted versus a varying ratio of beacon nodes (see Section 3.2 for further simulation details). We see a large fraction of nodes are unable to localize under the Centroid and Amorphous methods even with high beacon node ratios. Although in Figure 1 MCL has a low failure rate in providing location estimates, we observe these can be highly inaccurate with errors as high as four times the radio range. In the absence of application requirements commenting on such levels of error is difficult, however, [8] shows that geographic routing can perform adequately with errors of up to 0.4 times the radio range.

While existing mobile sensor deployments are small scale (e.g., ZebraNet [3] and Cartel [4] use fewer than ten nodes) and use GPS-equipped nodes with a relatively large form factor, in general, mobile sensor networks will include resource or size-constrained nodes (e.g., sensors embedded in

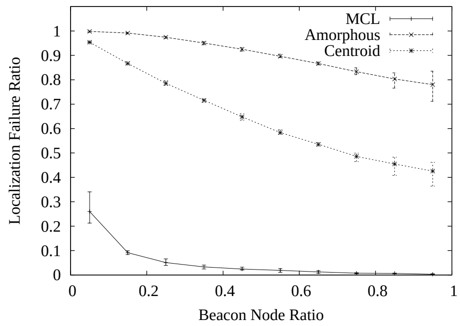

Figure 1: Comparison of a number of existing localization techniques, considering the ratio of failed localization attempts under variable beacon node density.

everyday mobile objects such as shoes [20], clothes, bikes, or attached to humans [2] or animals [3]). We are interested in targeting such embedded low-powered mobile sensors (e.g., mote class and smaller) that do not have access to GPS or a similar localization technique.

In this paper, we propose *ambient beacon localization (ABL)*, an alternative technique that generates location estimates by exploiting the sampled sensor data already requested by sensing applications. Our approach builds on the identification of distinguishing features reflected in the sensed data that can be mapped to physical locations. Some locations maybe easily recognized from the sensed data while many others may be too similar to be useful. We take a pragmatic approach to the problem and only attempt to recognize those sensed locations that are highly distinctive - we call these locations *ambient beacon points (ABPs)*. Under the ABL system, sensor data from the network is analyzed and once distinguishable features are identified they are mapped to location coordinates. This part of the system is currently based on an offline supervised learning stage. The learning stage produces a classifier algorithm that is able to map feature vectors to those coordinates that are recognizable. This classifier is disseminated to all mobile sensors. Nodes input their collected sensor data to the algorithm which outputs location co-ordinates when possible allowing the node to periodically self-localize. Because only distinct locations are used as ABPs they will tend to be only sparsely distributed across the sensor field. To increase the frequency at which mobile sensors can localize any sensor located at an ABP (therefore aware of its own position) assumes the additional role of a beacon node for the time it remains at the ABP. Mobile sensors receiving these ambient beacons can use a beacon-based localization algorithm (e.g., using [7]) to compute their own position estimate.

The benefit of ABL is two fold. First, the sensitivity of localization performance is no longer tied to connectivity between permanent beacons and non-beacons making ABL well-suited for mobile sensor networks. Second, localization leverages sensed data already collected by the main sensing function of the network. In a sense localization estimates *"come for free"* without the need for permanent beaconing infrastructure (such as GPS units) only requiring periodic retraining of the classifier. In this paper, we present the design of ABL and its initial evaluation in a building-sized testbed. While our research is at an early stage, initial experimental testbed and simulation results show the general utility of the idea. Additional work to fully analyze ABL is ongoing, using more extensive simulation, modeling, and large-scale experimentation.

2. AMBIENT BEACON DESIGN

The problem of identifying a set of unique features based on sensed data and binding these to all locations in a reliable and accurate manner is difficult, requiring sophisticated classification techniques, and addressing aliasing effects. Furthermore, the process must differentiate between the changes in observed sensor data that are due to the mobile node shifting physical position and those changes in the sensor data that are due to physical changes over time at the particular location itself. Thus, our approach is to *focus only on using locations whose characteristics produce relatively easily identifiable and largely time-invariant features in the sensed data.* Examples of such locations include those with: structures or obstacles (e.g., narrow alleyways, rock formations); unusual physical topology (e.g., speed bumps, stairs, wheel chair ramps, potholes); activities that are specific to the location (e.g., the rapid acceleration at a highway on ramp, animals drinking at a watering hole); predictable patterns in environmental data (e.g., a sharp rise in CO_2 and noise at the border of a city park, or a busy traffic intersection).

ABL uses supervised machine learning techniques to produce a simple classifier algorithm capable of mapping collected sensor data to a set of locations. Under supervised learning first a feature vector is defined, the elements of which are derived from the sensor data itself. Next, feature vectors that have been correctly labeled with locations are input to the training stage. Any suitable form of machine learning (i.e., such as decision trees, neural networks, belief networks or k nearest neighbors variants [9]) can be applied to these location labeled feature vectors. The output of the training stage is a classification algorithm that is disseminated to mobile nodes for later use in recognizing ABPs on the basis of gathered sensor data. The labels (locations) and feature vector (based on sensor data) are carefully selected resulting in a classifier of high accuracy with low state and computational requirements while requiring infrequent retraining. By only using a subset of locations that are most easily distinguished the accuracy and simplicity of the classifier can be maintained. The feature vector is only constructed from a subset of the elements found to be useful (e.g., the rate of change of temperature, the output of a compass, etc.), which are easily derived from the sensor data and largely time invariant. Such an approach can result in the reduction of possible ABPs across the sensor field but promotes a simplified classifier design and construction. By loosening these constraints (e.g., accepting less highly distinct ABPs) we can increase the ABP density across the sensor field but at the cost of increasing the complexity (e.g., requiring more frequent retraining) of the classifier.

In our ABL implementation, all nodes initially proceed through a bootstrapping phase where their classifier is initialized, after which it can be updated to reflect physical changes in the sensor field. Such updates are performed by retraining the classifier on the basis of periodically delivered sensor data that has already been location-stamped without the aid of the ambient beaconing process. This location-

	Period 1	Period 2	Period 3
Region 1	49 secs	71 secs	97 secs
Region 2	151 secs	197 secs	235 secs
Region 3	57 secs	105 secs	109 secs

(a) Interval between ambient beaconing assuming all building occupants during the experiment wore sensors.

	Period 1	Period 2	Period 3
Failure Ratio	0.07	0.19	0.26

(b) Ratio of nodes that failed to determine a location estimate during different observation period traces.

	Region 1	Region 2	Region 3	Default
Region 1	79%	4%	17%	0%
Region 2	0%	83%	11%	6%
Region 3	1%	15%	74%	10%
Default	1%	3%	6%	91%

(c) Confusion matrix for the location classifier.

Table 1: Results of Human Experiments

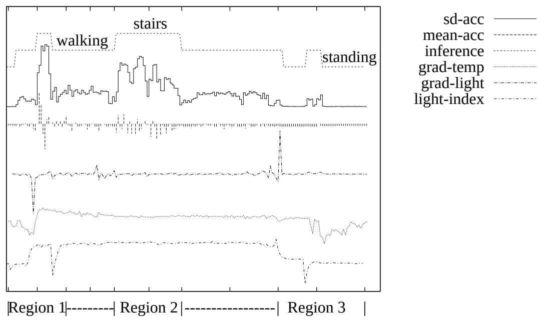

Figure 2: A representative time series trace from a sensor attached to an individual. The y-axis is a unitless multimodal representation of the elements of the feature vector. The x-axis provides the output classifier labels based on the feature vector.

stamped data is provided by independently location-aware mobile nodes (e.g., equipped with GPS) that either continuously roam the sensor field gathering data or are deployed only when classifier updates are necessary. We collect the training data and train the classifier as part of an offline process. The resulting classifier is then disseminated to mobile sensors from data collection points in the network.

3. EXPERIMENTATION

In what follows, we present a feasibility study of ABL including initial results from both a small scale human-based set of experiments, and simulations of a larger scale mobile sensor deployment. At this stage of our project we have not implemented a full ABL system but rather our experiments aim to answer questions concerning: (i) the feasibility of creating a location classifier; (ii) the impact of mobility patterns on ambient beacon generation; and finally (iii) the effectiveness of ambient beacons as a method of localization. We study these issues as an initial evaluation of the concept.

3.1 Human Experiments

For the human-based experiments, we consider a scenario where Moteiv Tmote Invent devices [10] are carried by all people in the Computer Science building at Dartmouth College. To approximate this scenario, we perform a number of experiments. In the first experiment ten people carry motes performing mobile sensing across the ground floor of the building ($\approx 1000m^2$). Motes sample and locally store temperature, light, and 2-D accelerometer data. Each sample is associated with a location region of $3m^2$ which we derive from manual measurements. The second experiment collects human pedestrian flow characteristics for occupants of the building. In the third experiments we extrapolate our findings across the entire building population.

We construct the classifier using the J.48 decision tree algorithm as implemented in Weka [9], a workbench that provides implementations of a variety of common machine learning algorithms. The output of the algorithm is a decision tree that classifies instances of feature vectors with a label associated with a location. In the experiment, the set of labels applied only covers three distinct locations with a fourth default label that represents all other locations. By keeping the set of labels small we look at a case where classifiable locations (ABPs) are sparse. We find under these conditions an effective feature vector includes the following: rate of change of light intensity; rate of temperature change; the raw temperature value; a moving mean and its standard deviation of the y-axis value of the 2-D accelerometer; the output of a three state (viz. walking, stationary, stair climbing) activity inference module based on clustering x-axis and y-axis accelerometer data; and an index indicating the ratio of artificial to natural light based on expected oscillation frequencies of florescent lights in this case. Figure 2 is an example time series trace showing the collected values of the feature vector changing due to the movement of an individual in the building. The values of the feature vector elements are shown on the unitless y-axis and are derived from the gathered sensor data. Along the x-axis of the plot we show the classifier output that results from this trace. Table 1(c) presents a standard confusion matrix representation of the classifier performance. In the table, the labels output by the classifier are shown as columns and the actual physical positions are shown as rows. The percentage in each table cell indicates the frequency over nineteen trials at which particular labels are assigned while the sensor is within one of the three regions/ABPs. Data used in the training phase is excluded from consideration. The classifier performs fairly uniformly across all three of the regions, generating the correct labeling approximately 80% of the time. The results of Table 1(c) are supported by 10-fold cross validation performed with the classifier against the training data; this test reported an accuracy rate of 83.3%.

The effectiveness of ABL is influenced by the arrival rate of mobile sensors at ABPs. Given the effect of deployment density we extrapolate our findings from only 10 mobile sensors, to the case where all building occupants are carrying mobile sensors. To do this we monitor the pedestrian flow in each of the regions that the classifier is trained to label. We use simple laser tripwire devices set up at each of the locations which transmit timestamped readings to nearby static sensors that are part of our permanent building sensor network testbed. We gathered traces during a one week period. From each trace we extract the average pedestrian

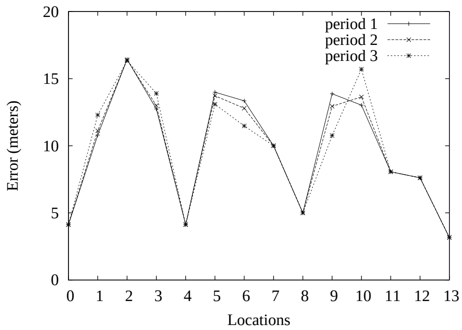

Figure 3: Localization accuracy achieved during the human experiments for a randomly selected set of positions within the building based on pedestrian traces taken at three different periods during a week.

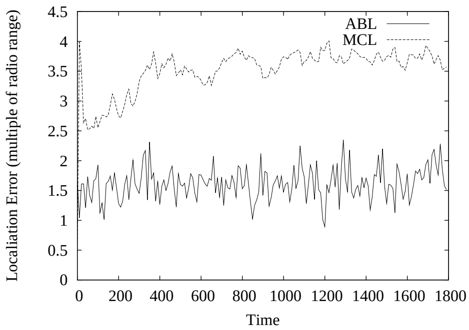

Figure 4: Comparison of the localization error of ABL relative to MCL.

arrival rate to each of the three classifiable locations, referring to them as periods one through three in our results. The mobile sensors beacon the location as long the classifier considers the sensor to be within the region. Since the tripwire only provides arrival times we combine these arrival times with the average beaconing duration at each of the three locations that we observe during the individual 10 person trials. Table 1(a) reports the interval between ambient beacons being emitted for each location using the combination of these arrival time averages and duration lengths.

To gage the efficacy of ambient beacons in determining location we place stationary Invent motes in each of the three classifiable locations and program them to beacon according the arrival rate during the pedestrian monitoring and for a beaconing duration based upon the 10 person experiments. This emulates the situation where all building occupants are carrying sensors that beacon while visiting a location they recognize (ABPs). At this point of in the evaluation five people wear mobile sensors and walk the same predefined path within the building. We measure the localization failure ratio and localization error for each person walking the path. The localization failure ratio is the number of times a location estimate cannot be computed either from ambient beacons or via the output of the classifier when localization is attempted every 15 seconds. Location estimates based on ambient beacons use the centroid-based technique from [7]. Localization error is represented as the difference between the location estimate and the actual location of 14 randomly selected distinct positions along the path. Each time period in which we monitor pedestrian traffic we observe slightly different arrival rates. Consequently, we perform all experiments using the differing arrival rates from each of these time periods. We repeat each experiment four times for each person, for each time period.

Table 1(b) reports the localization failure rate for each of the time periods given in Table 1(a). Figure 3 shows the localization error with the ambient beacon based estimates being in line with expectations of free-range beacon techniques under these conditions. Note, much of the error is due to difference between assumed and actual RF propagation.

3.2 Simulation Results

In what follows, we investigate the performance between a number of proposed localization schemes and ABL. In the scenario, all nodes are mobile but not all nodes are capable of independent self-localization. We use a discrete time event simulator with a pseudo-disk model for radio propagation, where a tunable random variable determines the amplitude of random variations over time around the nominal disk radius. The node mobility traces are based on actual GPS traces taken during the ZebraNet deployment [11]. Because the number of Zebras used in ZebraNet is small we augment these traces with semi-synthetic additional traces, applying the same methodology discussed in [12]. For all simulations we use 100 nodes moving in a sensor field sized 1000 by 1000 units. We present results based on ten experimental trials for each data point with each experiment spanning 1800 simulator time units. Nodes have a transmission range of 50 units.

We compare the performance of three existing techniques (viz. MCL [5], Amorphous Localization [6] and Centroid [7]), using implementations of these algorithms provided by the authors of [5]. At each discrete time step during the simulation localization is attempted for all nodes using each algorithm. Figure 1 shows the fraction of the localization attempts that fail as the ratio of beacon nodes in the network is varied while the size of the network remains constant.

To compare the performance of ABL to the other schemes under common conditions we simulate the operation of the ABL algorithm by applying observations from our human based experimentation. All simulated nodes are provided with an initialized classifier (learning has already occurred; we assume steady state performance w.r.t. classifier training). All nodes beacon on the basis of the classifier output, which has a simulated error frequency set according to Table 1(a). We set the density of ABPs to 5% of locations within the field. Classifiable locations are randomly selected until this ratio is achieved. The selection of these locations is randomly permuted for each experimental trial.

Figure 4 shows the location estimate accuracy versus elapsed simulation time for both the MCL technique (best among the three tested in Figure 1) and ABL. Error is defined as the distance between the estimated location relative to the actual one, normalized by the transmission radius. The results show that ABL outperform MCL. In other simulations we found the accuracy of ABL to be insensitive to the beacon ratio but sensitive to the density of discernibly unique locations in the field.

4. RELATED WORK

There is little work found in the literature associated with localization in mobile sensor networks. In [5], a beacon-based scheme is presented that modifies typical monte carlo localization (MCL) to use range free beaconing and to meet the resource constraints of typical sensor networks. In [15], the authors apply MCL to an approach that combines both ranging and range free forms of exchanges between nodes. In [16], the authors present an extension to support mobility using radio-interferometry. Mobility is discussed as a way to assist in the localization of static nodes in [13] [14].

Prior examples exist of machine learning been employed to assist in other aspects of the localization problem other than the one considered in ABL. In [17], a kernel based machine learning technique is applied to reduce the inaccuracies in range estimates in sensor networks, a building block of many localization schemes. [18] proposes Bayesian learning to build probabilistic localization filters used to localize robots, and uses both an awareness of the motion of the robot and the environment (i.e., such as doorways in buildings).

Finally, much can be learnt from the field of robotics when designing localization methods for mobile resource and size-constrained sensor networks. However, common assumptions about the environment (e.g., known layouts, artificial markers) and equipment (visual recognition systems [19][22], laser range finders) found in the robotics literature limits the direct application of these solutions.

5. CONCLUSION

In this paper, we presented ABL a novel approach for localizing mobile sensors based on leveraging ambient environmental sensor data. Sensed data is used to determine location and to a degree comes for free as part of supporting the sensing application. We presented some initial results that we intend to expand upon as part of future work. The results from experimentation and simulation show that the approach has promise. We are in the process of modeling a large mobile sensor network that considers the density functions for mobile sensors and their stochastic arrival processes.

Acknowledgment

This work is supported in part by Intel Corp., NSF NCS-0631289, and the Institute for Security Technology Studies (ISTS) at Dartmouth College. ISTS support is provided by the U.S. Department of Homeland Security under Grant Award Number 2006-CS-001-000001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

6. REFERENCES

[1] T. Abdelzaher, *et al.* Mobiscopes for Human Spaces In *IEEE Pervasive Computing*, 6(2), 2007.

[2] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, and R. A. Peterson. People-centric Urban Sensing. In *ACM/IEEE WICON 2006. Boston, MA, USA.*

[3] P. Zhang, C.M. Sadler, S.A. Lyon, and M. Martonosi. Hardware design experiences in zebranet. In *ACM SenSys 2004. Baltimore, MD, USA.*

[4] B. Hull, *et al.* CarTel: A Distributed Mobile Sensor Computing System. In *ACM SenSys 2006. Boulder, CO, USA.*

[5] L. Hu and D. Evans. Localization for mobile sensor networks. In *ACM MobiCom 2004. Philadelphia, PA, USA.*

[6] R. Nagpal, H.E. Shrobe, and J. Bachrach. Organizing a global coordinate system from local information on an ad hoc sensor network. In *IPSN 2003. Palo Alto, CA, USA.*

[7] N. Bulusu, J. Heidemann, and D. Estrin. Gps-less low cost outdoor localization for very small devices. *IEEE Personal Communications Magazine*, 7(5), 2000.

[8] T. He, *et al.* Range-free localization schemes for large scale sensor networks. In *ACM MobiCom 2003. San Diego, CA, USA.*

[9] I. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann, 2005.

[10] Moteiv Tmote Invent. http://www.moteiv.com/

[11] Y. Wang, *et al.* CRAWDAD data set princeton/zebranet (v. 2007-02-14). Downloaded from http://crawdad.cs.dartmouth.edu/princeton/zebranet, Feb 2007.

[12] Y. Wang, S. Jain, M. Martonosi, and K. Fall. Erasure-coding based routing for opportunistic networks. In *ACM SIGCOMM WDTN 2005. Philadelphia, Pennsylvania, USA.*

[13] P.N. Pathirana, N. Bulusu, A.V. Savkin, and S. Jha. Node localization using mobile robots in delay-tolerant sensor networks. *IEEE Trans. on Mobile Computing*, 4(3), 2005.

[14] N.B. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller. Mobile-Assisted Localization in Wireless Sensor Networks. In *IEEE INFOCOM 2005. Miami, FL, USA.*

[15] B. Dil, S. O. Dulman, and P. J. M. Havinga. Range-based localization in mobile sensor networks. In *EWSN 2006. Zurich, Switzerland.*

[16] B. Kusy, *et al.* intrack: High precision tracking of mobile sensor nodes. In *EWSN 2007. Delft, The Netherlands.*

[17] X. Nguyen, M. I. Jordan, and B. Sinopoli. A kernel-based learning approach to ad hoc sensor network localization. *ACM Trans. Sen. Netw.*, 1(1), 2005.

[18] S. Thrun. Bayesian landmark learning for mobile robot localization. *Mach. Learn.*, 33(1), 1998.

[19] R. Lerner, E. Rivlin, and I. Shimshoni. Landmark selection for task-oriented navigation. In *Proc. Of the IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, Oct 2006.

[20] Nike$^+$. http://www.nikeplus.com.

[21] Metrosense Project. http://metrosense.dartmouth.edu/.

[22] I. Ulrich and I.R. Nourbakhsh. Appearance-based place recognition for topological localization. In *IEEE Int'l Conf. on Robotics and Automation*, 2000.