# An Exploration of Location Error Estimation

David Dearman, Alex Varshavsky, Eyal de Lara, and Khai N. Truong

Department of Computer Science
University of Toronto
Toronto, Ontario M5S 3G4 Canada
{dearman,walex,delara,khai}@cs.toronto.edu

**Abstract.** Many existing localization systems generate location predictions, but fail to report how accurate the predictions are. This paper explores the effect of revealing the error of location predictions to the end-user in a location finding field study. We report findings obtained under four different error visualization conditions and show significant benefit in revealing the error of location predictions to the user in location finding tasks. We report the observed influences of error on participants' strategies for location finding. Additionally, given the observed benefit of a dynamic estimate of error, we design practical algorithms for estimating the error of a location prediction. Analysis of the algorithms shows a median estimation inaccuracy of up to $50m$ from the predicted location's true error.

## 1 Introduction

Developers of location-aware applications have a variety of localization systems [1,2,3,4,5,6,7,8] available to them. However, the accuracy and infrastructure requirements of each localization system constrains its applicability and the way in which application designers and end-users use them. For example, Wi-Fi [6] and GSM [7] localization are applicable for wide spectrum of applications, but their accuracy has shown to be highly variable. That is, the distance between the *actual location* and the *predicted location* fluctuates. We define this distance as the *true error* (see Figure 1) of the localization system. Unfortunately, many localization systems provide no information about the accuracy of their location predictions. As a result, the end-user of a location-aware application is entrusted to make sense of the location prediction presented to them without an awareness of the possible positioning error.

In this paper, we argue that presenting the error for a location prediction improves the usability of a location-aware application from the end-user's perspective. We present our examination of how users cope with an existing localization system and how they benefit from the presentation of the positioning error. We study user navigation strategies toward four predicted locations under four different error visualization conditions, where the user is presented with:

- only the predicted location and no additional information
- a region defined by a ring of fixed size, inside which the localization system is 95% confident that the actual location is contained

J. Krumm et al. (Eds.): UbiComp 2007, LNCS 4717, pp. 181–198, 2007.

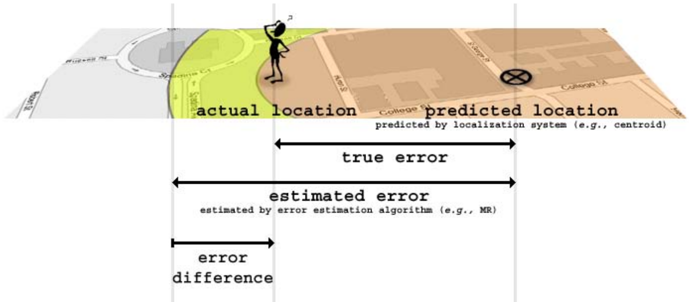


**Fig. 1.** Explanation of the positioning terminology used in this paper

– a region defined by a ring of variable size, inside which the localization system is N% confident that the actual location is contained; where N is defined by the user
– a region defined by a ring of variable size, inside which the localization system is optimally confident that the actual location is contained

In addition to understanding how users perform given the various error visualization conditions, we identify the effect of revealing the positioning error on users' navigation strategies for location finding. The results show that while a fixed estimate of positioning error provides little advantage to the user, a dynamic estimate of positioning error provides a significant benefit. A dynamic error estimate provides users with a better understanding of the true error at the time of prediction, whereas with a fixed error estimate, users are unable to differentiate between a high or low error location prediction until the actual location is found.

Finally, we describe two practical algorithms for dynamically estimating the error of a localization system and evaluate their performance. We show that our algorithms perform well, reaching a median difference of up to $50m$ between a location prediction's true error and the error as estimated by our algorithms.

## 2 Related Work

In this section, we introduce the most prevalent technologies used for location sensing. We discuss the importance of presenting their error for the usability of location-aware applications.

### 2.1 Location Sensing Technologies and Error

Many localization systems are available: Ultra-wideband[9], ultrasonic [1,2], infrared [3] , GSM [7], Wi-Fi [5,6], power lines [8] and GPS. The application of each system typically depends on its accuracy [10], infrastructure requirements and operating environment. Each of these systems has been experimentally tested [5,4,1,6,2] such that their error can be quantified for a particular environment. This measure is appropriate for validating the error of the localization system, but for end-users, it does not help them understand the true error of their *current* location's prediction. This issue is the focus of our research.

### 2.2 Expressing Uncertainty

Rather than refining the localization systems, our approach improves the usability of location-aware application from the end-user perspective. A location-aware application typically presents predicted locations to a user in such a way that it may require her to rationalize with the information [11]. Uncertainty in a context-aware application is inevitable [12], but the onus of identifying the true error of a predicted location should not be placed solely on the user. Rather, as Greenberg suggests, the context-aware application should not hide ambiguity and uncertainty from its users [13], but present the information truthfully in such a way that users can trust the information and react appropriately [14,15]. Chalmers and Galani advocate “seamful design”, in which systems designers reveal the finite nature or “seams” of their technology [16]. In doing so users and designers can leverage the system's finite nature (in our case positioning error) to provide a benefit. Antifaskos *et al.* [17] have shown in their memory aid study that user performance can be improved by expressing the system's uncertainty. Uncertainty is not always something that should be avoided. Depending on the context of use, uncertainty can be beneficial and used as a strategic element as seen in the mobile game ‘Can You See Me Now’ [15]. Participants were able to identify situations where GPS performed poorly and leverage the situations as part of their game play strategy. We believe that location-aware applications should truthfully present their location predictions as a measure of their *estimated error*, not simply the predicted location (see Figure 1).

Probabilistic localization systems that incorporate Kalman filtering or particle filters, typically have an internal representation of a confidence value in their location estimates that should be accessible by a system designed. GPS provides an easily accessible estimate of its uncertainty that is derived from the geometric dilution of precision (DOP). Of particular importance is the horizontal dilution of precision (HDOP). The HDOP does not provide a measure of error (*i.e.,* meters), but provides a scaling factor for the GPS receiver's accuracy based on the geometry of the visible satellites. A HDOP value of three or less indicates good satellite geometry and an accurate location estimate. Given the DOP, some GPS providers have implemented additional feedback that equates the DOP to an actual error measure.

## 3 Field Study Examining the Effects of Revealing Error of a Localization System

We conducted a between subjects field study that explores the effects of providing an estimate of the true error for a predicted location to users within the context of a location-aware map application. Specifically, we were interested in gauging the impact error has on users navigation strategies. To accomplish this issue, we conducted a location finding field study. Participants, aided by a location-aware mobile device, were required to find four posters positioned around the University of Toronto campus, while spending no more than 15 minutes to find each poster. We limited the time to find each poster to 15 minutes because it provided sufficient time to travel between poster locations (approximately three to five minutes) and search for the poster. Additionally, it limited the participant's exposure to the cold. We conducted the study during January and February of 2007 when the temperature around campus varied between -21°C and 1°C, while the weather varied from sunny to light snow.

### 3.1 Experimental Setup and Hardware

For the field study, we used Intel's POLS [4] GSM-based centroid algorithm as the underlying localization system for our location-aware map application. Our application and POLS were both installed on a Pocket PC T-Mobile MDA device. Prior to the study, we war-walked every street of the University of Toronto campus, carrying the T-Mobile MDA to train the centroid algorithm.

With the trained POLS system, we were able to analyze the true error of the centroid algorithm for our environment by identifying its cumulative distribution function. At the $95^{th}$ percentile, we determined the estimated error to be $467m$. Knowing the error levels associated with the positioning associated with our campus, we then chose four easily accessible locations that people frequently travel on campus such that the system would predict locations 1 and 3 with low error, and locations 2 and 4 with high error (see Table 1).

We physically marked the four locations with a unique poster (see Figure 2(a)). The posters came in two different dimensions: $61cm$(w) x $46cm$(h) and $40cm$(w) x $60cm$(h). Each poster was pressed against an equally sized piece of plywood and posted into the ground (see Figure 2(b)). None of the posters were placed inside a building, nor in an area that participants would have to walk through a building to find. The intent was to make each poster easily visible if the participants navigated within the poster's proximity. To motivate the participants to find each poster as fast as possible regardless of their previous failures or successes, for each poster, $25 was promised to the participant who found that location the fastest.

After installing a poster at each of the four locations, we used a T-Mobile MDA to collect 15 minutes of GSM measurements for each location and used the trained centroid algorithm to generate location predictions for each measurement. The location predictions for each poster were recorded in separate log files. The purpose of the logs was to ensure the consistency of the location predictions for all the posters across participants during the study. We then installed our map application on the same T-Mobile MDA. The application displayed a map of campus, annotated with the centroid algorithm's (live) prediction of the participants location and the predicted location of the poster (as read from the poster log file). Both locations were updated every two seconds. The map application provided three different zoom levels, where zooming in and out is achieved by pressing the respective up and down direction on the hardware directional pad.

**Table 1.** For each poster, the median, min, max and standard deviation of the centroid algorithms true error

| | Median | Minimum | Maximum | St.Dev |
|---|---|---|---|---|
| Poster 1 | 111$m$ | 13$m$ | 161$m$ | 40$m$ |
| Poster 2 | 228$m$ | 211$m$ | 237$m$ | 50$m$ |
| Poster 3 | 43$m$ | 19$m$ | 63$m$ | 11$m$ |
| Poster 4 | 248$m$ | 212 | 451$m$ | 39$m$ |


(a) Poster (b) Poster in context (poster circled)

**Fig. 2.** One of the four posters (a) placed around campus (b). Each poster displays a different image. The five digit code in the bottom right of the poster (a) is used by the application to validate the correct poster was found.

It is also possible to pan the map in any direction by applying pressure to the screen and dragging the map in the desired direction.

### 3.2 Participants

Thirty-two paid volunteers (27 male, 5 female) were recruited through the University of Toronto. Participants included both students and staff members from a variety of faculties. The age of participants was between 18 and 45, with most (27) between 20 and 35 years old. All participants are frequent computer users, interacting with a desktop and/or laptop computer on a daily basis. Most are mobile phone owners, but have varying experience with mobile computing beyond telecommunication (*i.e.,* text messaging or mobile WWW browsing): 14 are frequent (weekly) users, four are infrequent (monthly) users and 19 have no or little experience. Only three participants indicated previous experience with location-aware a technology (*e.g.,* GPS). Participants were drawn from the active university community to ensure they would be familiar with campus.

### 3.3 Experimental Conditions: Visualizing Error

We explored four technique of visualizing the error of the predicted locations: *predicted location, 95% confidence, customizable confidence* and *optimal error.*

**Table 2.** The mapping between the confidence level and the prediction error

| Confidence (%) | 5 | 10 | 15 | 20 | 25 | 30 | 35 | ... | 65 | 70 | 75 | 80 | 85 | 90 | 95 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Positioning error (m) | 57 | 82 | 103 | 123 | 139 | 153 | 164 | ... | 262 | 281 | 307 | 343 | 378 | 413 | 467 |

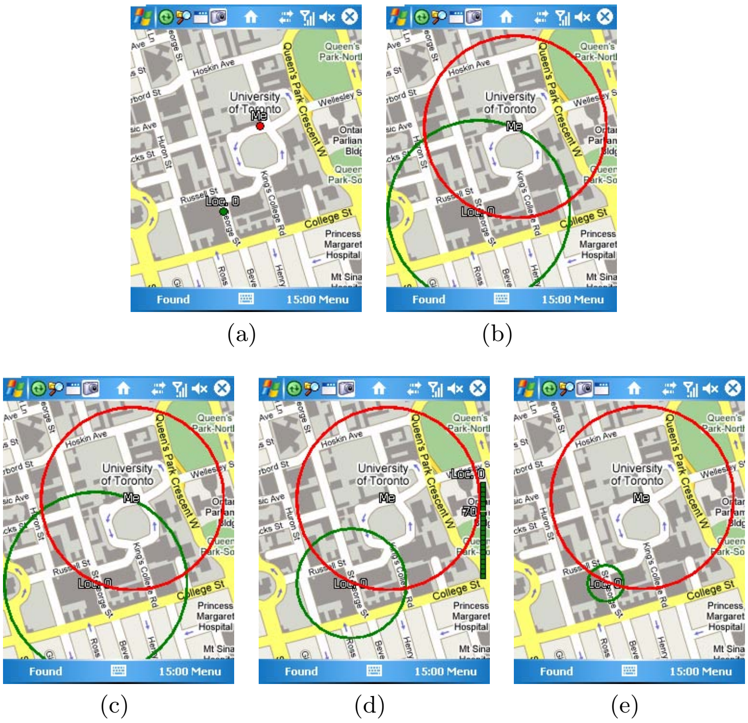


**Fig. 3.** The experimental four conditions: predicted location (a); 95% confidence (b); customizable confidence by default (c) and after the confidence value has been manipulated (d); and optimal error (e)

Given our between subjects study design, each participant was exposed to only one visualization, not multiple.

**Predicted Location.** In the predicted location condition, we provided participant with the predicted location of herself and the poster (see Figure 3(a)) as generated by POLS. This condition served as our base case to compare the other conditions against. The map is annotated with two dots representing the predicted locations of the participant and the poster. Participants were instructed that the error for the predicted locations could vary within the range of 467m.

**95% Confidence.** In the 95% confidence condition, we provided participants with a region defined by a confidence ring (see Figure 3(b)), in which the application is 95% confident that the actual location is contained within the ring.

The ring was drawn using our localization systems prediction of each location as the origin. The radius of the confidence ring was set to 467m; the $95^{th}$ percentile training error determined for our environment. The size of the confidence ring remained constant (in meters) throughout the experiment, but scaled (in pixels) to match the map's zoom level. The participants were not instructed that the confidence ring was drawn around an origin defined by the location prediction, but simply what the visualization represented.

**Customizable Confidence.** In the customizable confidence condition, we provided participants (by default) with the same visualization as the 95% confidence condition (see Figure 3(c)); however, they could manipulate the confidence level of the ring. By default the confidence is set to 95%, but by using the directional pad they could increase or decrease the confidence value respectively; the confidence is customizable in increments of five percent, from 5% to 95%. In changing the confidence value, the radius of the confidence ring would similarly change. A smaller confidence value would provide an smaller confidence ring; for example, Figure 3(d) shows the confidence ring set to 70%. If participants want a smaller area to search, they can decrease the confidence value, but in doing so they are decreasing the confidence the location is contained within the new confidence ring. Table 2 shows the relationship between the confidence levels and the ring size. Again, participants were not instructed that the confidence ring was drawn around an origin defined by the location prediction.

**Optimal Error.** In the optimal error condition, we provided participants with a ring for each location (see Figure 3(e)) where the ring's radius is defined by the true error of the location prediction. We could calculate the true error for each prediction because our software knew the actual location of both the participant (via GPS) and the posters. This means that the size of each ring is variable, but it provides optimal confidence because the rings always contain the actual location they represent. Again, participants were not instructed that the confidence ring was drawn around an origin defined by the location prediction. In actual practice, obtaining the true error of a location prediction given current technology is difficult at best. We attempt to address this issue in Section 4.

### 3.4 Procedure

The study began with participants filling out a background questionnaire to provide us with demographics information. Next, we introduced the participants to the experimental condition they would be using, explained the software and allowed them explore the interface. The explanation of the condition was repeated until the participants expressed an understanding of the condition. This ensured that the participants understood what the application was showing them. Each participant took part in only one experimental condition.

After the participants had sufficient time to explore the interface and were familiar with their condition, they began the actual experiment. We outfitted each participant with a voice recorder to record verbal comments made during the experiment and a Bluetooth GPS synchronized with the Pocket PC to record where participants walked. We escorted participants outside to the same initial location and instructed them to press the start button displayed on the handheld to begin. All participants were required to find the same posters, in the same order. No attempt was made to counter balance the poster ordering. Participants were instructed regarding a 15 minute time limit to find each poster. If after 15 minutes they were unsuccessful in finding the poster, the application indicated 15 minute time limit had elapsed and annotated the map the poster's actual location, to which the participant still must proceed. Once a participant had found a poster they entered the five digit code in the bottom right corner of the poster (see Figure 2(a)) to validate that they had indeed found the correct poster. If the poster code was valid, the application presented the participant with a Likert Scale questionnaire on the handheld inquiring about their perceived difficulty of the scenario. After completing the questionnaire, the application presented the participant with the map and the location for the next scenario. This process repeated for all four posters.

After finding the final poster and completing the questionnaire, the participants completed a semi-structured interview concerning their experience. Additionally, we asked them to rank the four posters according to difficulty.

## 4 Results of Field Study

In this section, we present the findings of our study. In particular, we focus on the participants' completion times and perceived difficulty for each poster. We highlight participants' navigation strategies and discuss the influence of the experimental conditions on their location finding strategies.

### 4.1 Time to Find a Poster

In Figure 4, we observe a significant dichotomy in the time to find a poster because of the magnitude of the estimated error; only 6 of 32 participants found poster 2 and only 10 found poster 4. Prior to our analysis, we applied the *Rankit*[1] procedure to normalize the timing data to make it more appropriate for variance analysis. The estimated error for a poster's location prediction and the condition had a significant affect on the time it took a participant to find a poster. A four (Condition) by four (Poster) analysis of variance (ANOVA) with the time to find a poster as the dependent variable, yielded a significant main effect for condition [$F(3,112) = 1.27$, $p<0.05$] and poster [$F(3,112) = 33.52$, $p<0.005$], but no significant interaction [$F(9,112) = 1.51$, $p>0.10$].

Post hoc comparisons were made using Tukeys HSD test. Results revealed that participants in the optimal error condition performed their poster finding task significantly faster ($p<0.05$) than those in the customizable condition. Additionally, those in the optimal error condition showed a trend towards performing the

[1] Timing data was normalized using the *Rankit* procedure as defined in SPSS v14.0.

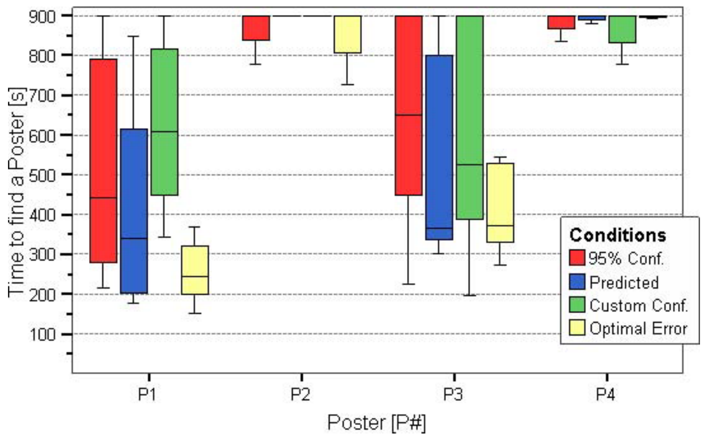


**Fig. 4.** Box plot of the time to find each poster for each condition. The box of the plot displays the median value and the interquartile range. The whiskers display the minimum and maximum time. All participants in the predicted and customizable confidence conditions could not find poster 2 in the 15 minute (900$s$) time limit.

poster finding tasks faster (p=0.06) than those in the 95% confidence condition. As expected, participants were able to find posters 1 and 3 significantly faster (p<0.001) than posters 2 and 4.

### 4.2 Perceived Difficulty of the Poster

The difficulty of each poster was assessed by two techniques. Upon finding a poster, participants rated the posters difficulty on a 5-point Likert scale (1-Very Easy, 2-Easy, 3-Neutral, 4-Difficult, 5-Very Difficult). After completing all posters, participants ranked the four posters in order according to their difficulty; 1 being the easiest and 4 the most difficult. The rank and Likert data is analyzed using Friedmans Two-Way Analysis of Variance. Post hoc analysis of the Likert and rank data is conducted using the Wilcoxon Signed-Ranks Test with a Bonferroni adjustment of $\alpha = 0.008$.

The Friedman analysis demonstrated a significant difference for the perceived difficulty between posters [$\chi^2$ (3, N = 32) = 35.58, p<0.001], a significant difference for the perceived difficulty between conditions [$\chi^2$ (3, N = 32) = 13.45, p<0.005] and a significant ordering of the posters perceived difficulty rankings [$\chi^2$ (3, N = 32) = 36.34, p<0.001].

Post hoc analysis of the conditions revealed participants perceived the optimal error condition to be significantly less difficult that the customizable confidence (Likert: z=-2.93, p<0.005) condition. Additionally, as expected given the estimated error, participants perceived poster 1 to be significantly less difficult than poster 2 (Likert: z=-411, p<0.001; rank: z=-4.72, p<0.001) and poster 4 (Likert: z=-3.40, p<0.005; rank: z=-3.72, p<0.001) and poster 3 significantly less difficult than poster 2 (Likert: $z=-3.74$, $p<0.001$; rank: $z=-3.55$, $p<0.001$) and poster 4 (Likert: $z=-3.37$, $p<0.005$).

### 4.3 Navigation Strategies

Each participant, upon finding all four posters created a retrospective route map by tracing their route from memory onto a paper map of campus. Using the map, participants could highlight unique occurrences during the experiment and easily convey locations alluded to during discussion with the interviewer. Using the retrospective map, a plot of each participants recorded position via GPS and their interview transcript, it was possible to identify typical and unique navigation strategies and the influence of each condition on these strategies. These strategies include:

- **Navigate to the middle.** In 99 of the 128 completed scenarios, participants either navigate to the predicted location and search the vicinity, or in the case of a ring, they navigate and search the region alluded to by the ring's centre. In the case of location predictions with low true error, this strategy was advantageous because it brought participants close to the posters; as such the majority of participants successfully find posters 1 (28/32) and 3 (24/32). However, for location predictions with high true error, navigating to the ring's centre region was detrimental in that it focused the participants search on an incorrect region; as such we observed only a small number of participants finding posters 2 (6/32) and 4 (10/32).
- **Confine the area to search.** In the predicted location condition, the application presented each poster as a single dot. Despite participants' understanding that error existed in the localization system (as described by the experimenter), participants' comments suggest that they struggled to translate the range of potential error into a meaningful search area. For location predictions with high true error, without an awareness of the error magnitude, participants typically searched one street (poster 4) or intersection (poster 2) exhaustively, without success (see Figure 4). For the most part, participants confined their search to too small an area and as such were unsuccessful in finding posters 2 and 4. In the three conditions in which participants were presented a region defined by the estimated error, the region helped to confine the participant's search. As mentioned, participants would typically start by searching the rings center region, but then expand their search to additional regions bounded by the estimated error ring.
- **Identify a path that provides the largest coverage of the surrounding area.** Rather than initially heading toward the error ring's centre region, some participants leveraged their understanding of the ring and their familiarity with campus to identify a path that would allow for maximum coverage of the suggested search area. This strategy included finding a sequence of streets and paths that would allow them to navigate the ring, encompassing as much area as possible, without backtracking.

- **Associate the target with a landmark.** Rather than initially going to the error ring's centre region, some participants relied on their knowledge of campus within the ring to specifically identify a unique location, or locations to search. They often associated a specific landmark or a well known location on campus as a probable location to find a poster. The justification for choosing the location(s) was based more on prior knowledge of the environment rather than the map. They made educated decisions as to where the poster could be based on their knowledge of the area within the ring.
- **Ignore the 'Me' location.** The majority of participants (27/32) explicitly expressed in their interviews that they ignored the application's prediction of their location for the majority of the study; they commented that their location prediction was not very accurate. However, 17 of the 32 participants indicated in their interviews that they did attempt to use their location's prediction on one or more occasions to help guide them. They tried to: 1) apply the error observed in the prediction of their location (given they knew where they were) to the predicted location of the poster, 2) align the 'Me' location with the poster location, and 3) infer greater accuracy when their predicted location and poster location overlapped. It was typical to see participants use these techniques with the first poster as their initial strategy (while they were inexperienced) or for the remaining three posters when they were unsuccessful in locating the poster where they expected. Most participants who relied on the 'Me' location at one point or another described their usage as a "last ditch" attempt to find the poster. Very few participants (3/32) repeated the same technique a second time for a subsequent poster. These strategies, although well conceived and seemingly plausible, were based on a nave understanding of the localization system; the perceived relationship between the two location predictions did not exist because the systems accuracy is variable depending on environmental features.

In addition to the strategies described above, we also identified important findings that highlight the need for presenting the estimated error of a localization system to users dynamically. In the customizable confidence condition, participants reduced the confidence ring, commenting in their interview, that is made the size of the prediction area more manageable. The application logs support this observation. This was appropriate for location predictions with low true error, however, it was detrimental in the case of high true error. A ring that dynamically shrinks and grows, such as the one in our optimal condition provides a more accurate awareness of the true error. Without this awareness, users are unable to differentiate between low and high true error.

- **Experience could not allude to the level of error.** The first two scenarios introduced participants to the fact that the error for a location predictions is variable. However, participants in the predicted location, $95^{th}$ percent confidence and customizable confidence condition did not show a significant change in their navigation strategies for posters 3 and 4. Participants commented that they searched regions closer to the rings edge (if they were presented with a ring), but overall most maintained the same strategy of searching the middle. Analysis of the GPS logs and the retrospective map supported their comments. Participant comments reveal that their strategy did not change because, 1) they had no awareness of the error level for the predicted location until they found the actual location of the poster and 2) they could not conceive a more beneficial strategy based on the information that they had:

> (P24-95th) *"In one [the first scenario] it [the poster] was in the centre, in two [the second scenario] it [the poster] wasn't. I didnt know what to choose! I chose the centre route [referring to scenario three] and I was lucky it was there."*

- **Desired reduction of the search area to a manageable size.** No participant in the customizable confidence condition maintained the 95% confidence level while trying to find the posters. All participants in this condition felt that the area defined by 95% confidence was too great to search within 15 minutes. At the default level of 95% confidence, the posters' actual location was always contained within the confidence ring. After participants reduced the confidence ring, the actual location was contained within the confidence ring 86% of the time for poster 1, 64% for poster 2, 99% for poster 3 and 43% for poster 4. For posters 2 and 4, given the high true error in the location prediction, reducing the confidence level resulted in the poster not being contained with the ring a substantial amount of the time.
- **Required awareness of true error.** Participants in the optimal condition were given a ring that provided them with an awareness of the true error for a location prediction. The size of the ring was proportional to the predicted locations true error: small error resulted in a small ring, large error resulted in a large ring. For posters 1 (M = $107m$) and 3 (M = $43m$) the size of the ring was significantly smaller than the ring for posters 2 (M = $228m$) and 4 (M = $266m$). For predictions with low true error, as in the case of posters 1 and 3, the participants had a much smaller region to search than the 95% confidence and customizable confidence conditions, but had the same confidence in the region. As such, we observed all participants in the optimal error condition found posters 1 and 3. However, only 12/16 found poster 1, and 10/16 found poster 3 for the 95% confidence and customizable confidence conditions. Additionally, the dynamic changes in the ring size afforded participants an awareness of the true error level for the current location prediction. As such, they could perceive the difficulty of each poster based on the size of the ring. In all the other conditions, participants did not have this awareness. They were ignorant of the true error for a prediction until they found the actual poster location.

> (P37-Optimal) *"[Referring to poster 2] I expected it to be more work, the circle was much larger than the first one."*

### 4.4 Summary

We observed a significant benefit of presenting the estimated error on participants ability to find the poster locations. The predicted location gave participants only one point to reference, as such they had difficulty defining an area to search when the predicted location was inaccurate. The 95% confidence, customizable confidence and optimal error conditions provided participants with a defined area to search. However, participants in the 95% confidence and customizable confidence condition could not identify the true error of the estimated error, as such they could not differentiate between an accurate and inaccurate location prediction. For the customizable confidence, participants found the default 95% presented them with an unmanageable search area. As a result, they reduced the confidence value which in the case of the low true error was beneficial, but in the case of high true error, it often resulted in the posters actual location being outside the confidence ring. Participants in the optimal condition had the benefit of a smaller ring provided the true error was low, but always had a consistently high level of confidence in their error estimates. We believe given our results that, not only is providing the localization error important, but that the presentation should present the true error as accurately as possible. In the next section, we address the issue of appropriately estimating the true error.

## 5 Dynamic Error Estimation

In the previous section, we showed that revealing the error of a localization system is beneficial to the end-user. In this section, we describe and evaluate two algorithms for dynamically estimating the true error of a location prediction as generated by a radio-based localization system such as centroid [10] or fingerprinting [7]. The algorithms are not tested in a similar field study as presented in the previous sections. We have left this exploration for future work.

### 5.1 Multiple Regression Error Estimation

The Multiple Regression (MR) error estimation algorithm takes as input a radio measurement (a list of beacons and their associated signal strength values), as fed into a localization system, and returns an error estimate of the location prediction in meters. To create a mapping from radio measurements to error estimations, MR uses the multiple regression method [18] to build a linear function from features of the radio measurements to error estimations. We experimented with a variety of features, eventually building a function that incorporates: the strongest signal strength value; the average of the three strongest signal strength values; the average of the five strongest signal strength values; the average of all signal strength values; the standard deviation of all signal strength values; the weakest signal strength value; the number of beacons observed; the number of strong signal strength values; the number of medium signal strength values; and the number of weak signal strength values. The strong, medium and weak signal strength values are device specific and need to be normalized across different devices. Once the linear function has been generated based on a set of training data, the error estimation can be generated in real time by extracting features from the current radio measurements and evaluating the linear function.

### 5.2 Zone Based Error Estimation

The Zone Based (ZB) error estimation algorithm takes as input a location prediction from a localization system (lat/lon coordinate) and returns an error estimate in meters. The main assumption behind the ZB algorithm is that localization systems are more or less stable. That is, if today a localization system predicts that the user is at coordinate B when she is actually at coordinate A, then tomorrow the localization system will still predict that the user is somewhere close to B, when she is at A.

ZB maintains a database of locations and errors associated with every location. Such a database is built off-line by running a given localization system on a set of measurements for which actual locations are known and recording the predicted locations and their associated error in the database. For example, predicted locations that fall around coordinate A may have a true error in the range of $100m$ to $120m$. This fact is recorded in the database.

Given the predicted location from a localization system, the estimated error is generated at real time by searching for known errors near the predicted location. Since a number of errors may have been recorded around the predicted location, ZB algorithm first sorts the errors and then chooses one of the errors based on an additional parameter given to the algorithm. For example, ZB50 uses the $50^{th}$ percentile (a median) value, while ZB75 uses the $75^{th}$ percentile value.

### 5.3 Evaluation

To evaluate the accuracy of the MR and ZB algorithms, we collected three sets of GPS-stamped GSM measurements using Intel's POLS software [4]. The data was collected by war-walking major streets of the University of Toronto campus. For each trace, we walked a distance of about $4.5km$, covering approximate area of $590,000m^2$. We evaluated MR and ZB on two localization systems: centroid and fingerprinting. We used the first set of collected data to train each localization system, and then tested their accuracy by feeding the two additional sets into each localization system to generate traces of GPS-stamped GSM measurements and corresponding location predictions.

Figures 5 and 6 show our algorithms performance for the centroid and fingerprinting localization systems, respectively. The figures show the $25^{th}$, $50^{th}$, $75^{th}$ and $95^{th}$ percentiles of the absolute difference between the true error and the estimated error for six error estimation algorithms. The Stat95 algorithm always predicts the same error, equal to the $95^{th}$ percentile of the error in the training data, while Random picks a random error estimation from the training data. Stat95 and Random are the straw man approaches and are presented for comparison. The numbers in parentheses represent the percentage of predictions

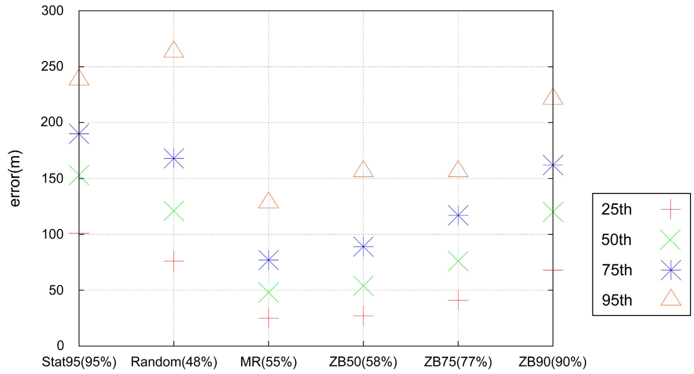


**Fig. 5.** The absolute difference between the true error and the estimated error for the centroid localization system using the Stat95, Random, MR and ZB algorithms

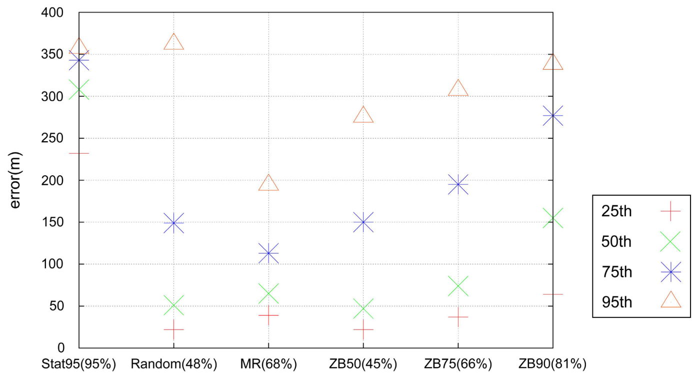


**Fig. 6.** The absolute difference between the true error and the estimated error for the fingerprinting localization system using the Stat95, Random, MR and ZB algorithms

for the respective algorithm that are greater than the true error. The percentage can be thought of as the confidence value for predictions generated by the algorithm. For example, an error estimate generated by ZB75 is approximately 77% more likely to be greater than the true error than lower.

Both MR and ZB perform better than the straw man approaches, with MR typically being more accurate. MR achieves $95^{th}$ percentile error of $128m$ for centroid and $194m$ for fingerprinting, with Stat95 (the best performing straw man) trailing behind with $238m$ for centroid and $357m$ for fingerprinting. ZB75 appears to have a good balance between accuracy and achieved confidence, but performs slightly worse in terms of accuracy than MR. Interestingly, there appears to be a high correlation between the parameters passed to the ZB algorithm and the achieved confidence level. This suggests that ZB may be used in systems where the confidence level may need to be adapted to the user's requirements.

**Table 3.** Median of the true error and estimation error (in meters) for the ZB75 and ZB90 algorithm at each of the four poster locations

| | Poster 1 | Poster 2 | Poster 3 | Poster 4 |
|---|---|---|---|---|
| True | 111 | 228 | 43 | 248 |
| MR | 196 | 208 | 142 | 124 |
| ZB75 | 122 | 196 | 266 | 298 |
| ZB90 | 148 | 243 | 519 | 364 |

As described in Section 3, we collected 15 minutes worth of location predictions at each for the four poster locations. To test the accuracy of our error estimation algorithms, we supplied these locations into our MR, ZB75 and ZB90 algorithms. Table 3 shows the median of the true error and estimation error for each poster. MR performs well for posters 1 and 2, but over estimates the error for poster 3 and under estimates for poster 4. Both ZB75 and ZB90 perform well for posters 1, 2 and 4, but they estimate a much larger error for poster 3. Looking more carefully at the data reveals that the training data around poster three contains a varied mixture of low and high error values. The nature of the ZB75 and ZB90 is to pick the $75^{th}$ and $90^{th}$ percentile error around the area of a prediction, therefore we observe high predictions of error for poster 3. We are developing techniques that can identify problematic areas such as poster 3 and reveal these inconsistencies to the user.

## 6 Conclusions and Future Work

Many localization systems exist and can be used in location-aware applications. However, the majority of these systems do not provide easy access to an estimation of the prediction error, if any at all. We introduced three techniques for presenting the estimated error to address this problem; *95% confidence*, *customizable confidence* and *optimal error*. We conducted a field study to explore the benefits and influences of presenting the estimated error on location finding, by comparing our three visualization techniques against simply presenting the predicted location. Our results show that presenting an estimate of the positioning error provides a significant benefit. Fixed estimates of error (*e.g.,* 95% confidence and customizable confidence) provided little additional benefit, but they do help confine the search area. The optimal error condition strongly and positively influenced participants' search strategies. Participants found all posters where the true error was small. When the true error was large, participants experienced the same problems for finding the posters as the participants in the other conditions. However, participants in the optimal condition could identify that the true error was large and differentiate between high and low true error, where as participants in all other conditions could not.

Based on the result of our field study, we designed two practical algorithms for estimating the error of a localization system. The Multiple Regression algorithm estimates the error based of the raw GSM measurements, by extracting features from the measurement and evaluating a linear function learned on the training data. The Zone Based algorithm generates an error estimate based on the predicted locations supplied by the localization system using a mapping of the predicted locations to errors. We evaluated the performance of our algorithms on the centroid and fingerprinting localization systems. Our algorithms perform well, showing a median estimation inaccuracy of up to $50m$ from the predicted location's true error.

In future work, we plan to continue with the refinement of our error estimation algorithms. The success of our algorithms and the simplicity of their design provide encouragement for the exploration of more robust methods. Additionally, we will explore alternate presentation of the estimated error beyond a simplistic ring. We believe that a different presentation may significantly improve the perception and understanding of the estimated error.

## Acknowledgments

The authors would like that thank all the participants for their time and willingness to brave the cold weather. This research is supported in part by the Natural Science and Engineering Research Council of Canada (NSERC) and the Walter C. Sumner Foundation.

## References

1. Harter, A., Hopper, A., Steggles, P., Ward, A., Webster, P.: The anatomy of a context-aware application. In: MobiCom '99: Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking, pp. 59–68. ACM Press, New York (1999)
2. Priyantha, N.B., Chakraborty, A., Balakrishnan, H.: The cricket location-support system. In: MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking, pp. 32–43. ACM Press, New York (2000)
3. Want, R., Hopper, A., Falcao, V., Gibbons, J.: The active badge location system. ACM Trans. Inf. Syst. 10(1), 91–102 (1992)
4. Chen, M.Y., Sohn, T., Chmelev, D., Hähnel, D., Hightower, J., Hughes, J., LaMarca, A., Potter, F., Smith, I.E., Varshavsky, A.: Practical metropolitan-scale positioning for gsm phones. In: UbiComp '06: Proceedings of the 8th international conference on Ubiquitous Computing, pp. 225–242. Springer, Heidelberg (2006)
5. Bahl, P., Padmanabhan, V.N.: Radar: an in-building rf-based user location and tracking system. In: INFOCOM '00: The 19th annual joint conference of the IEEE Computer and Communication Societies, pp. 775–784. IEEE Computer Society Press, Los Alamitos (2000)
6. Lamarca, A., Chawathe, Y., Consolvo, S., Hightower, J., Smith, I., Scott, J., Sohn, T., Howard, J., Hughes, J., Potter, F., Tabert, J., Powledge, P., Borriello, G., Schilit, B.: Place lab: Device positioning using radio beacons in the wild. In: Pervasive '05: The 3rd international conference on pervasive computing, pp. 116–133. Springer, Heidelberg (2005)
7. Otsason, V., Varshavsky, A., LaMarca, A., de Lara, E.: Accurate gsm indoor localization. In: UbiComp '05: Proceedings of the 7th International Conference on Ubiquitous Computing, pp. 141–158. Springer, Heidelberg (2005)

8. Patel, S.N., Truong, K.N., Abowd, G.D.: Powerline positioning: A practical sub-room-level indoor location system for domestic use. In: UbiComp '06: Proceedings of the 8th international conference on Ubiquitous Computing, pp. 441–458. Springer, Heidelberg (2006)
9. Ubisense, `www.ubisense.net`
10. Varshavsky, A., Chen, M., de Lara, E., Froehlich, J., Haehnel, D., Hightower, J., LaMarca, A., Potter, F., Sohn, T., Tang, K., Smith, a.I.: Are GSM phones THE solution for localization? In: IEEE Workshop on Mobile Computing Systems and Applications, April 2006, IEEE Computer Society Press, Los Alamitos (2006)
11. Smith, I., Consolvo, S., LaMarca, A., Hightower, J., Scott, J., Sohn, T., Hughes, J., Iachello, G., Abowd, G.D.: Social disclosure of place: From location technology to communication practices. In: Pervasive '05: The 3rd international conference on pervasive computing, pp. 134–151. Springer, Heidelberg (2005)
12. Dey, A., Mankoff, J., Abowd, G., Carter, S.: Distributed mediation of ambiguous context in aware environments. In: UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology, pp. 121–130. ACM Press, New York (2002)
13. Greenberg, S.: Context as a dynamic construct. Human-Computer Interaction 16, 257–268 (2001)
14. Edwards, W.K., Grinter, R.E.: At home with ubiquitous computing: Seven challenges. In: UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing, pp. 256–272. Springer, London (2001)
15. Benford, S., Anastasi, R., Flintham, M., Drozd, A., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., Row-Farr, J.: Coping with uncertainty in a location-based game. IEEE Pervasive Computing 02(3), 34–41 (2003)
16. Chalmers, M., Galani, A.: Seamful interweaving: heterogeneity in the theory and design of interactive systems. In: DIS '04: Proceedings of the 2004 conference on Designing interactive systems, pp. 243–252. ACM Press, New York (2004)
17. Antifakos, S., Schwaninger, A., Schiele, B.: Evaluating the effects of displaying uncertainty in context-aware applications. In: UbiComp '04: Proceedings of the 6th international conference on Ubiquitous Computing, pp. 54–69. Springer, Heidelberg (2004)
18. Krumm, J., Hinckley, K.: The nearme wireless proximity server. In: UbiComp '04: Proceedings of the 6th international conference on Ubiquitous Computing, September 2004, pp. 283–300. Springer, Heidelberg (2004)