RESEARCH   **Open Access**

Correlation-based radio localization in an indoor environment

Thomas Callaghan[1], Nicolai Czink[2*], Francesco Mani[3], Arogyaswami Paulraj[4] and George Papanicolaou[4]

Abstract

We investigate the feasibility of using correlation-based methods for estimating the spatial location of distributed receiving nodes in an indoor environment. Our algorithms do not assume any knowledge regarding the transmitter locations or the transmitted signal, but do assume that there are ambient signal sources whose location and properties are, however, not known. The motivation for this kind of node localization is to avoid interaction between nodes. It is most useful in non-line-of-sight propagation environments, where there is a lot of scattering. Correlation-based node localization is able to exploit an abundance of bandwidth of ambient signals, as well as the features of the scattering environment. The key idea is to compute pairwise cross correlations of the signals received at the nodes and use them to estimate the travel time between these nodes. By doing this for all pairs of receivers, we can construct an approximate map of their location using multidimensional scaling methods. We test this localization algorithm in a cubicle-style office environment based on both ray-tracing simulations, and measurement data from a radio measurement campaign using the Stanford broadband channel sounder. Contrary to what is seen in other applications of cross-correlation methods, the strongly scattering nature of the indoor environment complicates distance estimation. However, using statistical methods, the rich multipath environment can be turned partially into an advantage by enhancing ambient signal diversity and therefore making distance estimation more robust. The main result is that with our correlation-based statistical estimation procedure applied to the real data, assisted by multidimensional scaling, we were able to compute spatial antenna locations with an average error of about 2 m and pairwise distance estimates with an average error of 1.84 m. The theoretical resolution limit for the distance estimates is 1.25 m.

Keywords: indoor localization, sensor networks, signal correlations, rich scattering, multidimensional scaling

1 Introduction

Indoor localization is a long-standing open problem in wireless communications [1], particularly in wireless sensor networks [2,3]. Localization techniques in non-line-of-sight indoor environments face two major challenges: (i) multipath from rich scattering makes it difficult to identify the direct path, limiting the use of distance estimation based on time-delay-of-arrival (TDOA) methods; (ii) the strongly changing propagation loss due to shadowing impairs distance estimation based on the received signal strength (RSS).

In both kinds of algorithms, TDOA and RSS, nodes can estimate their own location relative to several "anchor nodes" acting as transmitters. This is commonly done by estimating the distances to the anchor nodes and subsequently using triangulation for position estimation.

The estimation of the TDOA is done either by round-trip time estimation [4], the transmission of specific training sequences [5], or simply by detecting the first peak of the received signal [6]. Ultra-wide band communications are specifically suited for TDOA distance estimation because of the large available bandwidth [7].

Many publications discuss RSS-based distance estimation. The work presented in [8] provides a comprehensive overview of an actual implementation using WiFi hotspots in a self-configuring network.

Another technique described in [9] uses spatial signatures for localization. However, this requires multiple antennas at the nodes and a database of spatial locations.

* Correspondence: czink@ftw.at
[2]FTW Forschungszentrum Telekommunikation Wien, Austria
Full list of author information is available at the end of the article

Moreover, this technique is limited to specific antenna requirements.

Correlation-based methods [10] have been widely used in the last few years in a variety of fields, including sensor networks. Some examples include estimation of the local propagation speed of surface seismic waves and even earthquake prediction [11]. The idea is to cross correlate seismic noise signals from seismographs deployed in a wide area so as to estimate the travel time of the seismic waves from one sensor to the other. Given the sensor locations, the wave speed can be estimated using travel time tomography.

Contribution

In this paper, we investigate the feasibility of *passive, correlation-based indoor radio localization*.

In contrast to previous works, our localization scheme only relies on *ambient signals with wide bandwidth*. Thus, *no dedicated transmitters* need to be deployed as long as the ambient signals from other wireless systems are sufficiently rich. In effect, the *radio signals are unknown*, the *location of the sources is unknown*, even *the number of effective sources is unknown*.

Even under these very stringent conditions, the distances between the receiving nodes can be estimated in a three-step procedure: (i) first, all nodes are receiving and recording ambient signals, (ii) the nodes communicate their received signals to a central entity or node, (iii) the central entity estimates the pairwise travel times, hence the distances, between all the nodes by cross correlating their received signals and identifying peaks in the cross-correlation function. If the ambient signals have sufficient spatial diversity, then the peaks of the cross correlations provide a robust estimate of the distance between the two receiving antennas. By doing this for all pairs of receiving nodes, we construct an approximate map of their locations using weighted least-squares methods, in particular multidimensional scaling (MDS) [12,13].

This method suggests that there are several *advantages for radio localization*:

• There is no communication overhead between nodes by active probing. Ranging is done *without nodes cooperating or even communicating* with each other. Nodes do not even know how many other peers are in their vicinity.
• Only the central entity has the information from which to estimate the location of the nodes. The gains of cooperative localization (i.e., the pairwise distance estimates between peers) are achieved at the central entity, without having the nodes cooperate. This is advantageous for situations, where nodes do not want to reveal their location to other peers, as with active probing.

• While the performance of TDOA ranging methods is inherently limited by the bandwidth of the (known) *transmitted* signals, correlation-based localization is only limited by the bandwidth of the (unknown) *received* signals, depending on communication or other wireless activities in the environment. Thus, correlation-based methods are not limited by scarce bandwidth allocations. Using wide-band receivers, a much higher ranging resolution can be obtained by simply recording *ambient signals* from any occupied bands. By that, the performance improves with the employed bandwidth of the receivers.

To show the feasibility of this approach, we explore the performance of correlation-based radio localization in an indoor environment. To quantify it, we use (i) ray-tracing simulations and (ii) data from a recently conducted radio measurement campaign, using the RUSK Stanford multi-antenna radio channel sounder with a center frequency of 2.45 GHz and bandwidth of 240 MHz [14].

The strongly scattering nature of the indoor environment makes the pairwise distance or travel time estimation challenging. However, in contrast to other localization methods, multipath from rich scattering is now both helpful and harmful for distance estimation. While multipath increases spatial diversity of the signals, it also leads to additional peaks in the correlation function that reduce the robustness of travel time estimation. The main feature in this work is the proper treatment and utilization of the beneficial properties of rich multipath while controlling its negative effects. To achieve this goal, we propose statistical peak-selection algorithms that significantly increase the localization accuracy.

We demonstrate, therefore, that passive, correlation-based radio localization is feasible in wireless indoor environments.

Organization

The paper is organized as follows. Section 2 provides a brief motivation for using correlation-based methods for distance estimation. In Section 3, we consider the problem of travel time estimation using cross correlations. Section 4 presents different approaches for improving the pairwise travel-time estimation based on correlation-based methods. Section 5 briefly presents how we use MDS to find position estimates, discusses the results from applying our algorithms and MDS to the simulated and measured data, and demonstrates the effect of transmitter positions using the simulated data. With Section 7, we conclude the paper. Appendices A.1 and A.2 provide brief descriptions of the ray-tracing simulations and the measurement data we use in this paper.

2 Motivation for the Use of Cross Correlations in Distance Estimation

We start out with a simple example. Consider a line-of-sight environment, as shown in Figure 1. A single source emits a pulse $s(\tau) = \delta(\tau)$, while two receivers record the signals $r_1(\tau)$ and $r_2(\tau)$, respectively, where τ denotes the delay and $\delta(\cdot)$ denotes the Dirac delta function. The positions of the source and of the receivers are unknown. The signal emitted by the source is received by both receivers with certain delay lags. Thus, the received signals become $r_1(\tau) = \gamma_1 \delta(\tau - \tau_1)$, and $r_2(\tau) = \gamma_2 \delta(\tau - \tau_2)$, where τ_k denotes the delay lag from the source to the kth receiver, and γ_k denotes the path loss of the signal. By cross correlating the two received signals,

$$c_{1,2}(\tau) = \int r_1(\tau')r_2(\tau + \tau')\mathrm{d}\tau' = \gamma_1\gamma_2\delta(\tau - (\tau_1 - \tau_2)),$$

we see that the resulting cross correlation is a pulse at the delay difference $\Delta\tau = \tau_1 - \tau_2$. This also holds for arbitrary source signals, as long as they have certain autocorrelation properties, as shown in the next section.

By finding the peak in the received signals cross correlation, we can estimate the distance between the receivers as $\hat{d} = \Delta\tau\, c_0$, with c_0 indicating the speed of light. When the transmitter is on a straight line going through the two receivers, this estimated distance is the exact distance between the nodes [10]. However, when there is an angle α between the direction of the plane wave front and the straight line between the receivers, the distance estimate will give $\hat{d} = d \mid \cos(\alpha) \mid$, which carries a systematic error.

Since we do not know the position of the source, we cannot correct for this systematic error, but we can quantify its distribution. For this, we make the following assumptions: (i) we consider horizontal wave propagation only, since it is predominant in indoor environments; (ii) all directions of the transmitted signals are equally likely, i.e., α is distributed uniformly, $\alpha \sim \mathcal{U}[-\pi, \pi]$). So, we can calculate the probability density function of the estimated distance, $p_{\hat{d}}(\hat{d})$ by

transformation of the random variable α as

$$p_{\hat{d}}(\hat{d}) = \begin{cases} \dfrac{2}{\pi} \dfrac{1}{\sqrt[d]{1 - (\hat{d}/d)^2}} & \dots \ 0 \le \hat{d} \le d, \\ 0 & \dots \ \hat{d} > d, \end{cases} \tag{2}$$

and also obtain its cumulative distribution function

$$F(\hat{d}) = \frac{2}{\pi} \arcsin\left(\frac{\hat{d}}{d}\right) \dots 0 \le \hat{d} \le d, \tag{3}$$

which is shown in Figure 2. It turns out that in 50% of all cases, our distance estimation error is less than 30% (indicated by the dashed lines).

While basing the distance estimation on a single plane wave is questionable because of the rather large systematic error, real radio propagation environments provide directional diversity by multiple sources and by multipath.

Multipath is both advantageous and challenging: (i) The receiver cross correlation gets multiple peaks providing more information about the propagation environment, which improves distance estimation, (ii) By reflections, the length of some paths can actually exceed the distance between the nodes.

Note that in this scheme, the existence of a direct line of sight (LOS) or non-LOS *between the nodes* is of reduced interest. More important is whether a wave can travel unobstructed over a pair of nodes. While we may observe an obstructed direct LOS between the nodes, we may still get a good distance estimate from another wave front connecting the node pair from a different propagation angle.

The way to exploit this signal diversity and how to obtain a robust distance estimate is the topic of the rest of this paper.

3 Computation of Cross Correlations

This section describes the computation of the cross correlations using a more complex setting with multiple sources, including scattering in the environment. A finite number of L sources, S_l, $l = 1,...,L$, transmit random uncorrelated signals $s_l(t, \tau)$, i.e.,
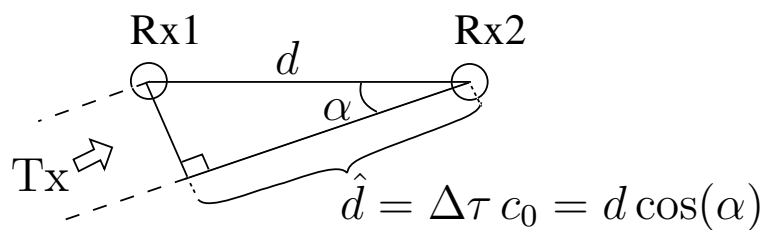

Figure 1 A plane wave from a single source is observed with a specific delay at both receivers. The delay difference is used to estimate the distance between the receivers
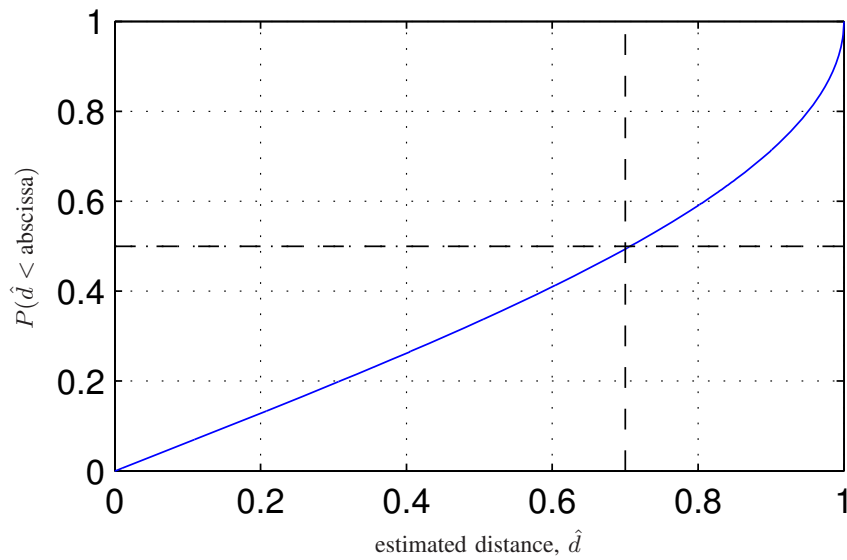
Figure 2 Cumulative distribution function of $\hat{d}$ for $d = 1$, assuming a uniform distribution of the direction of the impinging wave. In 50% of the cases, the estimation error is less than 30%

$$E\{s_l(t, \tau)s_{l'}(t', \tau')\} = \begin{cases} 1 \ldots l = l' \wedge \tau = \tau' \\ 0 \ldots l \neq l' \vee \tau \neq \tau', \end{cases} \quad (4)$$

where $\wedge$ denotes the logic AND operator, $\vee$ the logic OR operator, t denotes absolute time (assuming block transmission), and τ denotes the delay lag. For example, white noise signals fulfill these properties asymptotically, when $\tau \to \infty$. We assume that the channel stays constant within the transmission of a block and then changes due to fading. A number of K receivers, R_k, $k = 1,...,K$, record their respective received signals $r_k(t, \tau)$ from these multiple random sources, i.e.,

$$r_k(t, \tau) = \sum_{l=1}^{L} \int_{\tau'} s_l(t, \tau') h_{kl}(t, \tau - \tau') d\tau', \quad (5)$$

where $h_{kl}(t, \tau)$ denotes the time and frequency selective radio channel from the lth source to the kth receiver.

The cross-correlation function (CCF) between two received signals at time t is

$$c_{k,k'}(t, \tau) = \int_{\tau'} r_k(t, \tau') r_{k'}(t, \tau + \tau') d\tau', \quad (6)$$

which can be written as

$$c_{k,k'}(t, \tau) = \sum_{l=1}^{L} \int_{\tau'} h_{kl}(t, \tau') h_{k'l}(t, \tau + \tau') d\tau', \quad (7)$$

when the source signals fulfill the condition in (4). This CCF provides information about the delay lag

between the two receivers R_k and $R_{k'}$ as discussed in the previous section.

When applying this method to radio channel measurements, the CCF can be averaged over all measured time instants T (i.e., averaging over fading variations of the channel) by

$$\hat{c}_{k,k'}(\tau) = \frac{1}{T} \sum_{t=1}^{T} c_{k,k'}(t, \tau). \quad (8)$$

For the actual implementation, all convolutions and correlations in delay domain are implemented as multiplications in frequency domain.

It is well known [10] that for an infinite number of (uncorrelated) orthogonal sources, isotropically distributed in space, the resulting CCF has a rectangular shape, centered at zero and having a width of $2d/c_0$. The range resolution is limited by the bandwidth of the source signal and is given by c_0/B [15] due to using peak-search in a signal of limited bandwidth. In our setup, $c_0/B = 1.25$ m. Since in our simulations and measurements (cf. Appendices A.1 and A.2) only a finite number of transmitting antennas contribute to the signal recorded at each receiving antenna, we rely on sufficient scattering in the environment for enhancement of directional diversity. This leads to a trade-off between two effects: (i) Multipath increases the signal diversity and thus creates peaks in the CCF that better represent the true distance, but (ii) multipath also generates "wrong" (additional) peaks from propagation paths that do not directly travel through the receivers, which in turn reduce the accuracy of distance estimation.

An example of signals received at a pair of receivers and a CCF evaluated from our measurements (cf. Appendix A.2) is shown in Figure 3. We observed a strong directionality of the impinging radio waves, which leads to peaks at various distances. The true distance of 4.9 m, indicated by the dashed lines, is clearly visible as a peak in the CCF. However, other strong peaks are also present. Because of these multiple peaks, which sometimes dwarf the accurate peaks, a more elaborate distance estimation method is necessary.

4 Improved Distance Estimation Method

The distance estimation can be improved by combining four ideas: (i) using short-time estimates of the CCF, (ii) using multiple peaks from the CCF for distance estimation, (iii) using relative weighting on the peaks from the CCF to distinguish between peaks of comparable height (power), and (iv) using multi-dimensional scaling (MDS) to jointly improve the distance estimation and produce a location estimate.

As explained in detail above, given sufficient source diversity and a weakly scattering environment, the peak of the cross-correlation of signals recorded by two sensors in the environment corresponds to the travel time between them. However, little-to-no theory exists for the case of limited source diversity and a strongly scattering environment. In this situation, we have multiple strong peaks where possibly none correspond to the correct travel time. As a result, we developed an empirical approach to peak selection that tries to utilize the

information we have from both multiple peaks in the correlation functions and multiple realizations of the multipath in the environment. Others have studied how to address multiple peaks in cross-correlation in reverberant environments and developed strategies using secondary peaks, weighting, and a type of fourth order correlation function [16,17]. Peak selection in an optimal way is a challenging problem that will be the subject of future work.

4.1 Short-time Estimates of the CCF
The long-time averaging applied in the original approach in (8) may reduce information about the propagation environment. By using the short-time estimates of the CCF from (6), individual differences in the propagation environment, caused by fading, can be utilized to improve the distance estimation as follows.

4.2 Multiple Peaks for Distance Estimation
As motivated in Section 2, the distance between two receivers is proportional to the propagation delay,

$$\hat{d}_{k,k'} = \hat{\tau}_{k,k'} \, c_0, \qquad (9)$$

where $\hat{d}_{k,k'}$ and $\hat{\tau}_{k,k'}$ denote the distance estimate and travel time estimate, respectively.

A direct way to estimate the delay between two receivers is to identify the largest peak in their CCF. This approach does not perform well in multipath environments. Instead, we consider a more robust statistical approach based on *multiple peaks* in the CCF. The
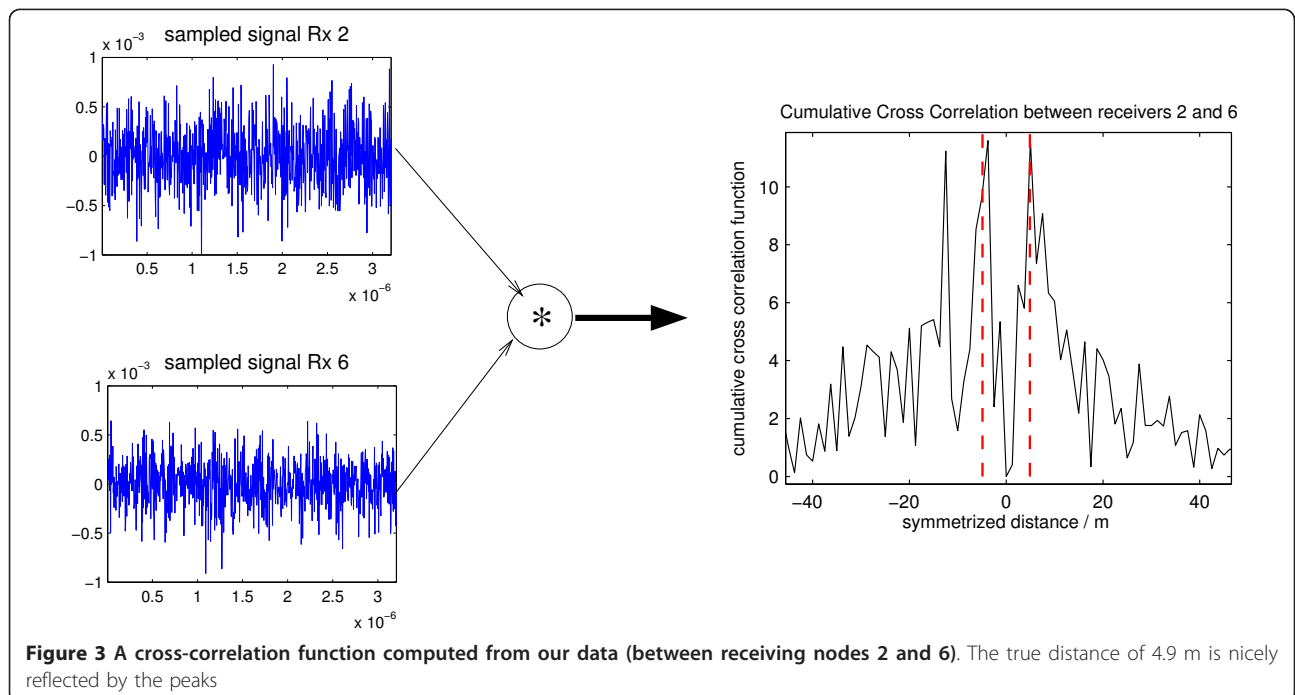

Figure 3 A cross-correlation function computed from our data (between receiving nodes 2 and 6). The true distance of 4.9 m is nicely reflected by the peaks

problem is how to choose and how to use the peaks in the CCF.

We use a statistical approach as follows: the CCF is sorted according to

$$\hat{c}_{k,k'}(t, \tau_1) > \hat{c}_{k,k'}(t, \tau_2) > \cdots > \hat{c}_{k,k'}(t, \tau_M), \quad (10)$$

with M denoting the number of resolved delays in the CCF. From this sorting, we use $p = 0.5\%$ of the delays having the strongest CCF values, i.e.,

$$\hat{\tau}_{k,k'}(t, n) =| \tau_n |, n \in \left[1, \dots \lfloor pM \rfloor\right], \quad (11)$$

which corresponds to taking the top $\lfloor pM \rfloor = 4$ peaks in our data set. The value of p should balance the trade-off between choosing enough peaks to average both the under- and over-estimation of the travel time and choosing few enough to exclude peaks that do not add useful information to travel-time estimation. Our choice for p is based on our empirical observations of the data.

We then take a weighted average of these multiple delays as the distance estimate, i.e.,

$$\hat{\tau}_{k,k'} = \frac{1}{T \cdot \lfloor pM \rfloor} \sum_{t=1}^{T} \sum_{n=1}^{\lfloor pM \rfloor} w_{k,k'}(t, n)\hat{\tau}_{k,k'}(t, n), \quad (12)$$

where the choice of the weights $w_{k,k'}(t, n)$ is described in the next section.

4.3 Cross-correlations with Weights

To improve the distance estimation further, we propose to distinguish between dominant peaks and peaks of similar amplitude. For this reason, we weigh the peaks based on their relative amplitude.

Since we are using $N = \lfloor pM \rfloor$ peaks, we assigned a weight to each peak equal to the ratio between its amplitude over the Nth largest peak's amplitude,

$$w_{k,k'}(t, n) = \hat{c}_{k,k'}(t, \tau_N)/\hat{c}_{k,k'}(t, \tau_N), n \in [1, N]. \quad (13)$$

The estimates computed by this statistical procedure can subsequently be improved by taking geometrical considerations into account as shown in the next section.

4.4 Multidimensional Scaling

Multidimensional scaling (MDS) algorithms are statistical techniques dating back 50 years, that take as its input a set of pairwise similarities and assign them locations in space [12,13]. Recently, it was applied to a different, but related problem, of node localization in sensor networks [3].

In our problem, the input is the distance estimates between all receiver pairs. Multidimensional scaling, after introducing a few more assumptions as stated

below, improves these individual distance estimates by jointly estimating receiver positions. The estimated receiver positions are also of much interest in this problem and are not simply a by-product in improving pairwise distance estimates.

In MDS, we have the following least-squares optimization problem

$$\min_{\{R_k\}} \sum_{k \neq k'} \lambda_{k,k'} | \hat{d}_{k,k'} - \| R_k - R_{k'} \|_2 |^2, \quad (14)$$

where $\hat{d}_{k,k'}$ are the provided distance estimates, $\lambda_{k,k'}$ are weights, $|| \cdot ||_2$ is the Euclidean distance, and R_k is the location of receiver k. In our problem, we assume the locations lie in $\mathbb{R}^2$. As we will show in the next section, the error in our pairwise distance estimates is correlated with the distance estimates themselves. Thus, a natural weighting is $\lambda_{k,k'} = 1/\hat{d}_{k,k'}^{\alpha}$, for $\alpha \geq 0$. We found that $\alpha = 1$ produced the smallest mean squared localization error. To solve this optimization problem, we used the algorithm given in [18], which resulted in the final position estimates $\{R_k\}$ of the receivers. The results using this algorithm are sensitive to the initial guess, so we used the following procedure to compute our initial position estimate:

1. To fix our initial receiver location, we first choose the receiving antenna $R_{k(1)}$ that has the smallest average estimated distance from the other receiving antennas and place it at the origin, i.e., $R_{k(1)} = (0,0)$.
2. The second receiver $R_{k(2)}$ is then chosen to be the one with the smallest estimated distance from the first receiver and is placed at $R_{k(2)} = \left(\hat{d}_{k(1)k(2)}, 0\right)$.
3. The third receiver $R_{k(3)}$ is then chosen to be the one with the smallest estimated distance from receivers $R_{k(1)}$ and $R_{k(2)}$ and placed at the point in the first quadrant $\hat{d}_{k1k3}$ from $R_{k(1)}$ and $\hat{d}_{k(2)k(3)}$ from $R_{k(2)}$. Should the third receiver fall on a line with the first two anchors, the triangle inequality is not valid and the space not properly spanned. In this case, another third receiver is chosen.
4. The rest of the receiving antennas are placed using the iterative least-squares lateration procedure in [19].

With position estimates computed using MDS, which jointly uses the pairwise distance estimates, we can compute new pairwise distance estimates. These distance estimates should be an improvement as they are "jointly computed" and explicitly use the geometry of the setup, i.e., the receivers lie in a 2-D plane.

5 Results

In the subsections that follow, we apply these distance estimation and localization methods to both a simulated dataset and data from an indoor radio channel measurement campaign.

5.1 Ray-tracing Simulations

We first present our localization method applied to a environment simulated using a state-of-the-art ray-tracing tool including diffuse scattering. A detailed description of the ray-tracing algorithm is provided in Appendix A.1. The nodes were set up as shown in Figure 4. We *estimated the distance* between all pairs of nodes and calculated the distance estimation error using the ("statistical peak selection") method presented in Section 4. As reference to with which to compare, we use the long-time average peak method ("cumulative peak"), i.e., selecting the strongest peak out of the averaged long-time CCF given in (8).

A scatter plot of the true distance versus the estimated distance for these approaches is shown in Figure 5. This plot exhibits a significant underestimation bias. We expect this from the theory, especially when we have strong sources illuminating from the "wrong" angle.

The empirical cdfs of the distance estimation errors are shown by the dashed lines in Figure 6. Using only the peak of the averaged long-time CCF performs worst, by far, because it does not use the diversity in time. In contrast, using our statistical peak selections significantly lowers the distance estimation error.

With the simulated channel bandwidth of 240 MHz, our theoretical resolution is limited to an accuracy of $c/B = 1.25$ m. Our final results produced an average pairwise distance estimation error of 4.55 m.

Next, we used MDS to obtain position estimates. By the weights introduced in the MDS, [a] we make use of the correlation between the distance estimation and its error. In Figure 7, the localization results using our statistical method are shown. The true locations are denoted by circles, while the estimated locations are marked by squares. The arrows are connecting the estimates to their respective true locations. The positions computed in the MDS need to be anchored by a frame of reference as translation, rotation, and reflection in the 2-D plane do not affect them. Three anchor positions would be needed to anchor the entire network. Instead of visualization, we find the rotation, translation, and reflection that gives the closest positions to the true locations in the least-squares sense.

Looking at this figure, we notice that the error is mostly in the *x*-direction. The reason for this is the strong directionality of the waves coming mostly from top/bottom, but not from left/right. This naturally leads to an underestimation of the distance between the horizontally-spaced node pairs. We also observe that the receiving antennas that are lying centrally have the smallest position estimation errors. This is due to the increased diversity of the source signals. We find an average position estimation error of 3.66 m, with a minimum error of 1.25 m, a maximum error of 5.87 m, and a standard deviation of 1.56 m.
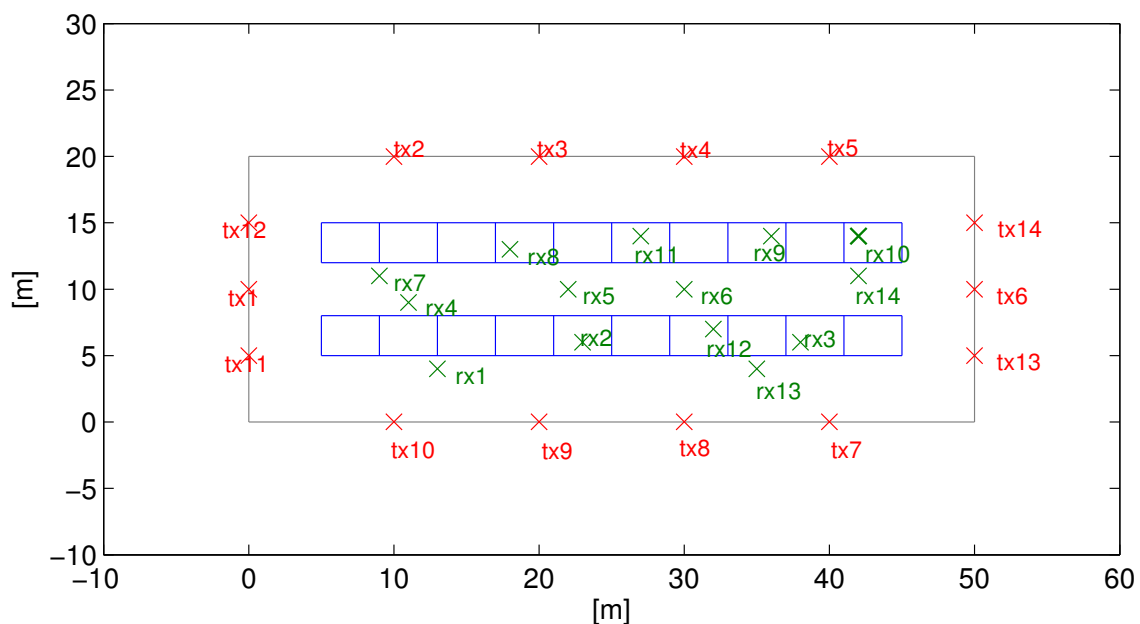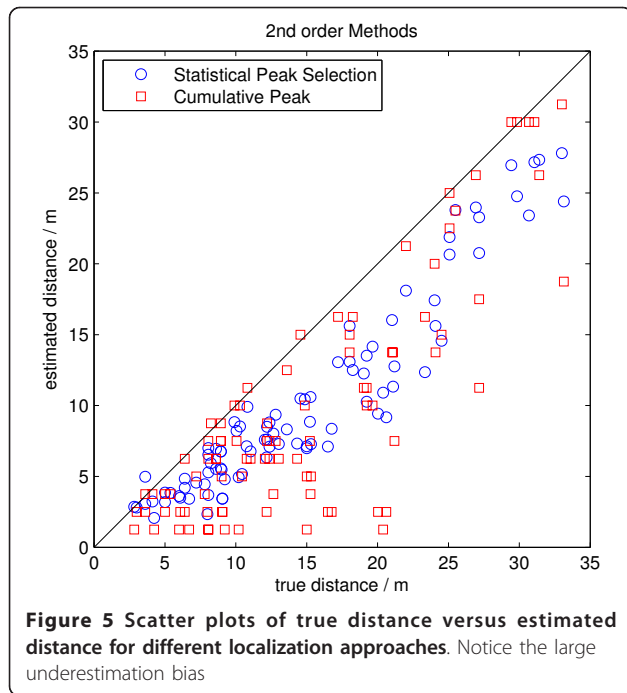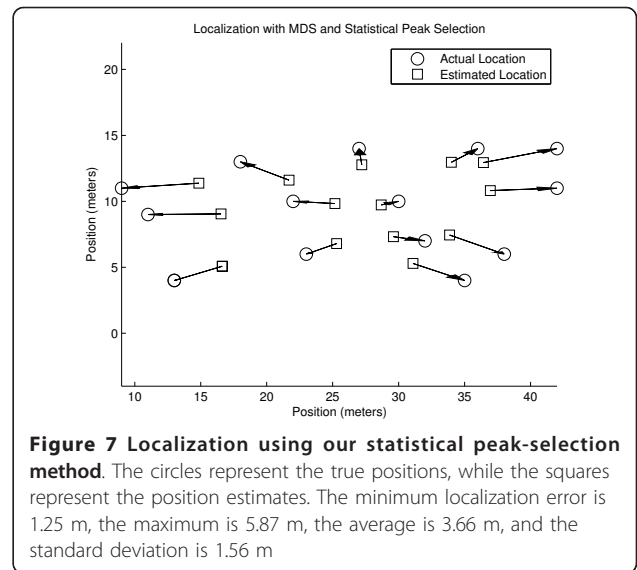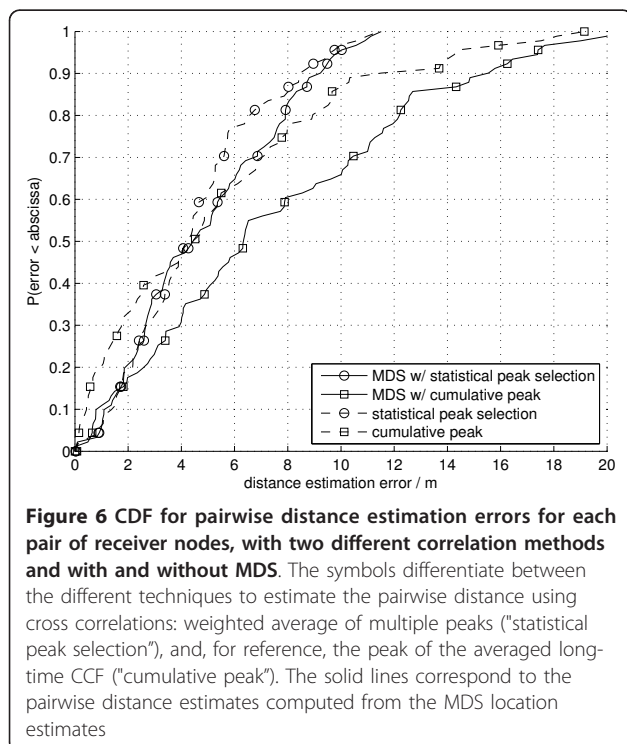

Figure 4 Location of transmitters and receivers for simulations. The red *x*s are the transmitters around the perimeter, and the green *x*s inside are the passive receivers. There are 14 of each. The scale is in meters

Figure 5 Scatter plots of true distance versus estimated distance for different localization approaches. Notice the large underestimation bias

Looking at the position estimates when using only the long-time average peak in Figure 8, the results seem questionable. Some distances are strongly underestimated as already seen in Figure 5. In this approach, reliable position estimation becomes impossible. This clearly demonstrates that multipath must not be

Figure 6 CDF for pairwise distance estimation errors for each pair of receiver nodes, with two different correlation methods and with and without MDS. The symbols differentiate between the different techniques to estimate the pairwise distance using cross correlations: weighted average of multiple peaks ("statistical peak selection"), and, for reference, the peak of the averaged long-time CCF ("cumulative peak"). The solid lines correspond to the pairwise distance estimates computed from the MDS location estimates

Figure 7 Localization using our statistical peak-selection method. The circles represent the true positions, while the squares represent the position estimates. The minimum localization error is 1.25 m, the maximum is 5.87 m, the average is 3.66 m, and the standard deviation is 1.56 m

ignored, but needs to be utilized to enable acceptable distance estimation.

To compare our position estimates to a benchmark, we compute the localization Cramer-Rao lower bound as in [20]. Limits of cooperative localization have also been studied in [21]. The distance estimation error is modeled by

$$\sigma_{ij}^2 = K_E d_{ij}^{\beta}, \tag{15}$$

where K_E is an environment factor, d denotes the distance between two nodes, and β an appropriate exponent. After calibrating this model with our simulations and using the equations from [20], we can quantify the CRLB of the estimation error for every individual node. The results are summarized in Table 1.

If our estimator is optimal (fulfilling the CRLB), then the mean value of the last column should be 1. In our
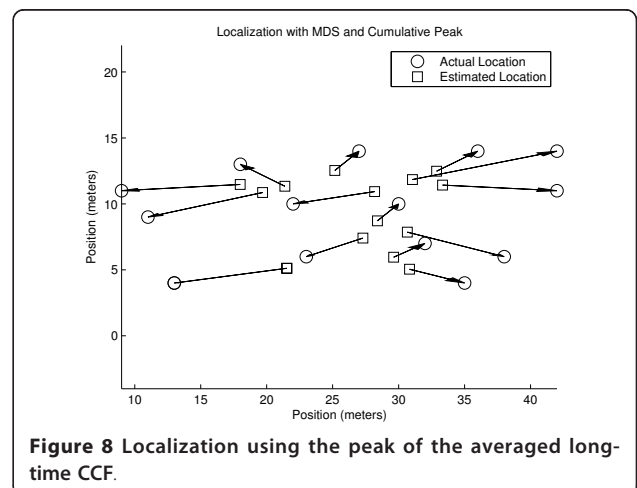

Figure 8 Localization using the peak of the averaged long-time CCF.

Table 1 CRLB versus localization errors

	CRLB	$3\sqrt{\text{CRLB}}$	$|R_k - \hat{R}_k|$	$\frac{|R_k - \hat{R}_k|^2}{CRLB_k}$
Simulations	1.74 m^2	3.92 m	3.66 m	9.69
Experiments	0.51 m^2	2.09 m	2.10 m	10.03

case, the mean value is ≈10. In other words, the variance of our distance estimator is about 10 times higher than the one of the CRLB; however, this estimate is based on just 14 samples.

Additionally, we can recompute pairwise distance estimates from the position estimates. The empirical cdfs of the distance estimation error are shown by the solid lines in Figure 6. While we expect these new distance estimates to be improved because they are computed jointly with the other receiver pairs constrained to lie on the plane via MDS, we see that this is not the case with the simulations. Since the distance estimates in our simulations are almost always underestimated, the MDS fails to improve over the initial distance estimates and rather makes the whole "environment" smaller.

5.2 Effect of Transmitter Locations

As mentioned earlier, the locations of the transmitters have a great effect on the quality of the distance and position estimation. With our simulated data, we can actually turn off and on certain transmitters and examine the effect that this has. Figure 9 demonstrates that selecting different sets of transmitters produce very different results. These figures use four different sets of transmitters: all, top and bottom, left and right, and the configuration that gave the minimal average position error in a thorough but not exhaustive search.

As expected, using the top and bottom transmitters results in good location estimation in the y-direction while using the left and right transmitters gives good location estimation in the x-direction. Comparing the location estimates of the top left and bottom right scatterers in Figure 9d to their position estimates using all of the transmitters (plot (a) in the same figure), one can observe that including the transmitter closest to the true receiver location results in that receiver's estimated position error being larger. This is also consistent with the intuition brought forward in Section 2. Sources close to the receiver nodes will most likely lead to an underestimation of the distance.

5.3 Performance in a Measured Environment

As a proof of concept, we applied our localization method to an indoor radio channel measurement described in Appendix A.2 The nodes were set up in two squares as shown in Figure 10. As with the ray-tracing simulations, we estimated the distance between all pairs of nodes using the ("statistical peak selection")
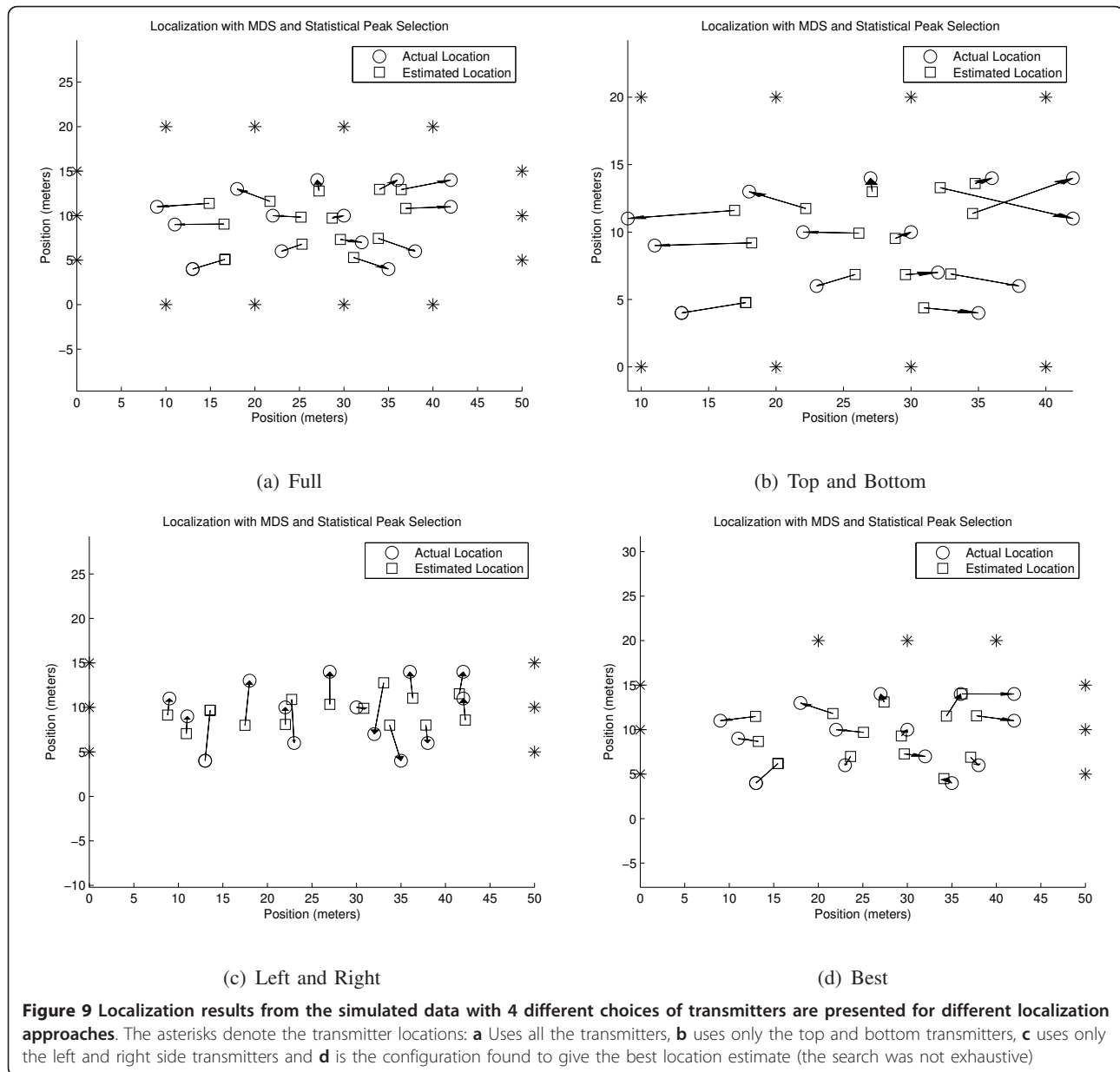
method presented in Section 4. As reference to with which to compare, we use the long-time average peak method ("cumulative peak"), i.e., selecting the strongest peak out of the averaged long-time CCF given in (8).

A scatter plot of the true distance versus the estimated distance for these approaches is shown in Figure 11. The interesting fact noted here is that for larger true distances, the distance estimation error becomes larger. This effect can be easily explained by the underlying wave propagation: our approach needs strong waves traveling through the receiver pair. When the receivers are far apart, the probability of a direct wave from one to the other becomes much lower. This is also the reason why the long-time average peak method performs so badly. The distance between the nodes is strongly underestimated. Only when making use of fading, i.e., diversity in time, the distance estimates become reliable. It is important to note that this result is significantly different than the result with the simulated data. This is due to the fact that the real measurements capture more of the complexity of the rich-scattering channel, and also contain measurement noise.

Again, the empirical cdfs of the distance estimation errors are shown by the dashed lines in Figure 12. With the measurement bandwidth of 240 MHz, our theoretical resolution is limited to an accuracy of c/B = 1.25 m. Our results produced an average pairwise distance estimation error of 2.33 m. Moreover, the distance estimation errors of almost half of our 28 receiving antenna pairs were below the resolution limit, which is again an effect of using the diversity offered by the time variations in the channel. Furthermore, when we recompute pairwise distance estimates from the position estimates, we see that joint estimation using MDS improves the results. This is shown by the solid lines in Figure 12. The average pairwise distance error is then 1.84 m.

We also present the results of the *location estimation* in Figure 13. The true locations are denoted by circles, while the estimated locations are marked by squares. The arrows are connecting the estimates to their respective true locations.

Looking at the quadrangle of the bottom four nodes, we observe that the estimates are placed in a rhomboid. The reason for this is the strong directionality of the waves coming mostly from left/right, but not from top/bottom. This naturally leads to an underestimation of the distance between the vertically-spaced node pairs. The result is that the nodes appear squeezed in the y-direction, but do have the correct distance in the x-direction. We also observe that the receiving antennas that are lying centrally have the smallest position estimation errors. This is due to the increased diversity of the source signals. We find an average position estimation error of 2.1 m, with a minimum error of 0.4 m, a

Figure 9 Localization results from the simulated data with 4 different choices of transmitters are presented for different localization approaches. The asterisks denote the transmitter locations: **a** Uses all the transmitters, **b** uses only the top and bottom transmitters, **c** uses only the left and right side transmitters and **d** is the configuration found to give the best location estimate (the search was not exhaustive)

maximum error of 3.36 m, and a standard deviation of 0.92 m.

Again, we compare these results to a benchmark using the Cramer-Rao lower bound. The results are summarized in Table 1. The variance of our distance estimator is about 10 times higher than the one of the CRLB; however, this estimate is based on just 8 samples. For our estimation scheme, this is quite a good result, leading to useful estimates in indoor environments. Note that even though the direct LOS between some nodes is sometimes obstructed by people, the distance estimation is still reasonable. This is due to wave fronts from other directions, which are not obstructed. Thus, our

algorithm is inherently robust against NLOS problems, as long as wave fronts can propagate over both nodes in a non-obstructed way.

As before, when comparing our method to using only the long-time average peak in Figure 14, the results are inaccurate and unreasonable.

6 Implementation and Complexity
A realistic implementation of these methods would of course require the consideration of several practical issues, including timing synchronization and information exchange between the receiver node and the central entity, and optimal selection of the radio band for

Figure 10 Floor plan of the cubicle-style office environment used for the localization measure-ments. R × 1-R × 8 are the locations of the receiving nodes while T × 1-T × 8 are the locations of the antennas generating the ambient signals
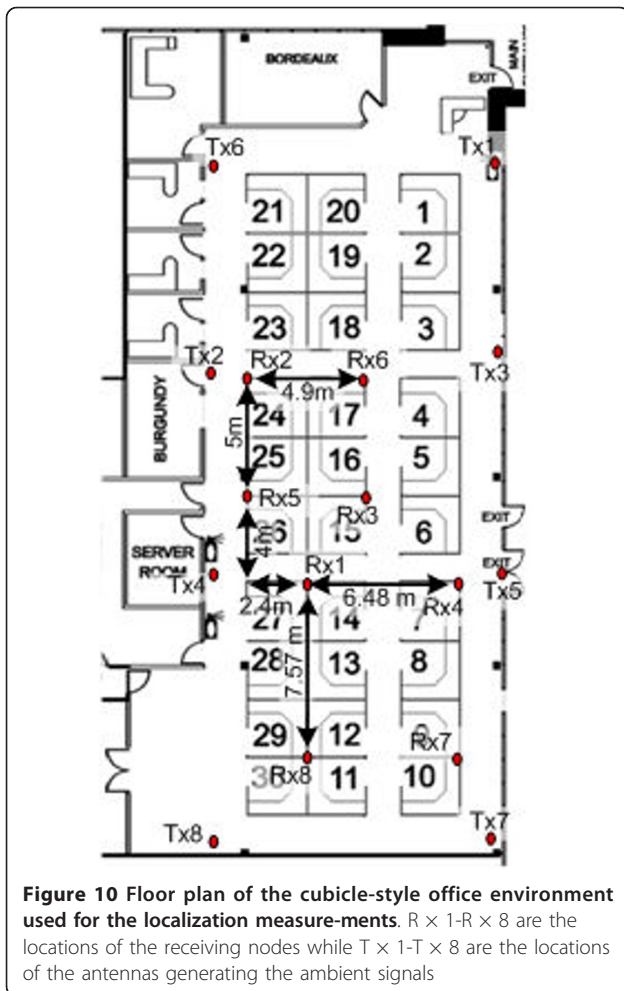
providing enough ambient signal strength. Also note that the central entity can be connected to the nodes by any means of wired or wireless communication.

Since the central entity performs all the calculations, there is no communication or ranging overhead between the nodes. The central entity can also ensure the synchronization between the nodes [22].

As for all delay-based localization algorithms, the receiving nodes need to sample the ambient signals with a high sampling rate (and thus bandwidth) using a fast analog-to-digital converter. The advantage of our approach is that the sampling can be done with a low bit resolution, against which our approach is robust. Of course, the recorded data can be further compressed before sending it on to the master.

As numerical example, we consider a bandwidth of $B = 240$ MHz, and a maximum resolvable distance of $d_{max} = 50$ m, a minimum number of $2Bd/c$ samples must be acquired. An ADC resolution of 8 bit is more than sufficient for the correlation-based ranging, thus the sampled data to be transmitted equals to 80 bytes per recorded

block. When employing data compression algorithms, this number is reduced even further (which is most beneficent when recording multiple blocks). Thus, only a limited amount of data needs to be transmitted, leading to a much smaller necessary communication bandwidth than what was needed for sampling the ambient signals.

The complexity of the localization problem can be segmented into the complexity for ranging, and the complexity for estimating the node positions. Since the latter complexity is the same for both kinds of algorithms, we just compare the ranging complexity.

The complexity of our scheme per distance estimation sample for K nodes is comprised of K signal sampling, storing, and communication events (nodes transmitting the recorded signals to the central entity), and $K(K-1)/2$ computations of the signal cross correlations and peak searches. Note that the computation of cross correlations can be done computationally efficiently using the fast Fourier transform. In contrast, conventional TDOA-based schemes need $K(K-1)/2$ ranging actions (i.e., ranging between all pairs of nodes). Subsequently, the ranging information must be communicated by at least $K-1$ nodes. Thus, the main difference in complexity lies in the computation of the cross correlations.

The complexity increase of the proposed scheme needs to be seen together with the advantage of the algorithm's robustness against NLOS environments, strongly reduced communication and coordination overhead, and the much higher available bandwidth and thus ranging resolution.

7 Conclusions

In this paper, we consider the feasibility of radio localization in a rich-scattering indoor environment using correlation-based techniques, where nodes only use unknown ambient signals for localization and do not probe actively.

We presented a systematic way to use peaks in the cross correlations of the received signals for computing pairwise distance estimates and spatial location estimates for a passive network of wireless receiving nodes (sensors). The robustness of the estimation is enhanced by multipath due to scattering but its accuracy is diminished by it. The increased signal diversity improves the estimation robustness while generating many peaks in the cross correlations. To enhance inter receiver (sensor) distance estimation, we use statistical methods that exploit multipath effects by taking into account multiple fading realizations of the channel.

The advantages of this scheme are threefold: (i) there is no ranging or communication overhead between the nodes, all communications and calculations are done by a central entity; (ii) the algorithm is inherently robust
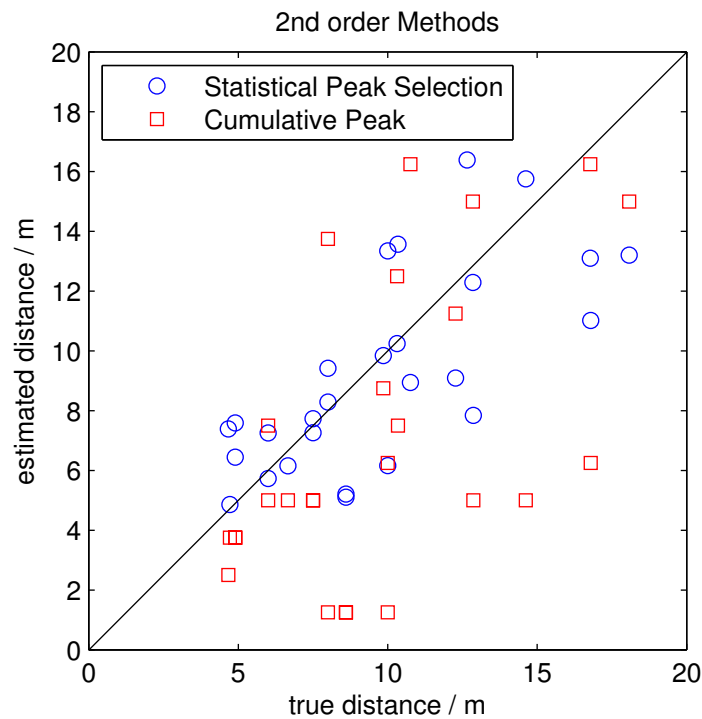
Figure 11 Scatter plots of true distance versus estimated distance for different localization approaches.

against non-line-of-sight between nodes; (iii) by using just unknown ambient signals for localization, there are no regulatory limitations of bandwidth (and thus resolution).

We demonstrated the feasibility of our approach using both simulated and real measurements in a cubicle-style office environment. In our simulations, we use a 3-D ray-tracing tool, operating at 2.45 GHz, to measure the radio
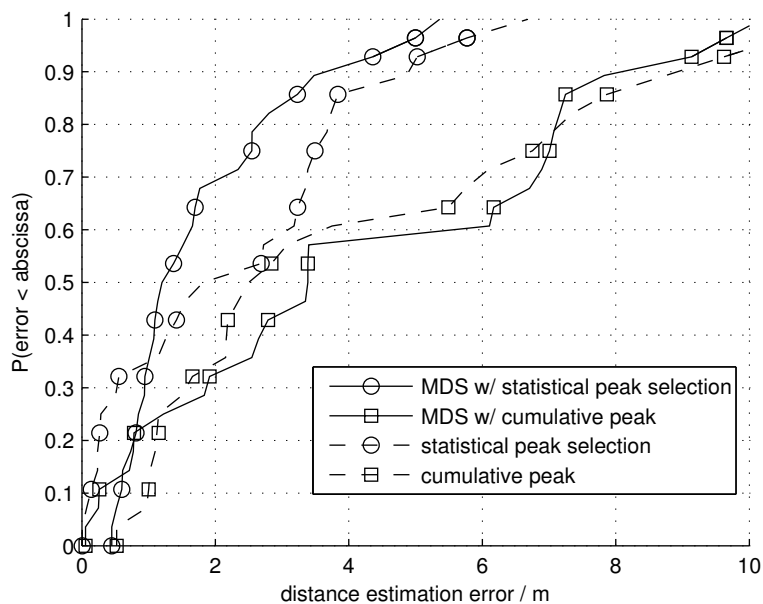
Figure 12 CDF for pairwise distance estimation errors for each pair of receiver nodes, with two different correlation methods and with and without MDS. The symbols differentiate between the different techniques to estimate the pairwise distance using cross correlations: weighted average of multiple peaks ("statistical peak selection"), and, for reference, the peak of the averaged long-time CCF ("cumulative peak"). The solid lines correspond to the pairwise distance estimates computed from the MDS location estimates
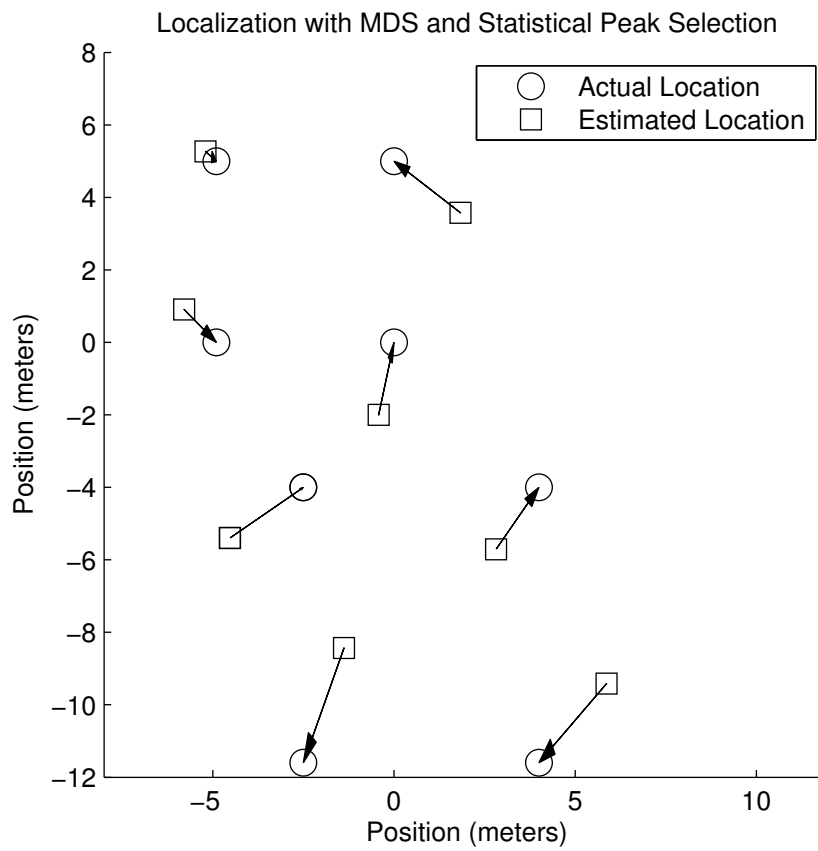
Figure 13 Localization using our statistical peak-selection method. The circles represent the true positions, while the squares represent the position estimates. The minimum localization error is 0.40 m, the maximum is 3.36 m, the average is 2.10 m, and the standard deviation is 0.92 m
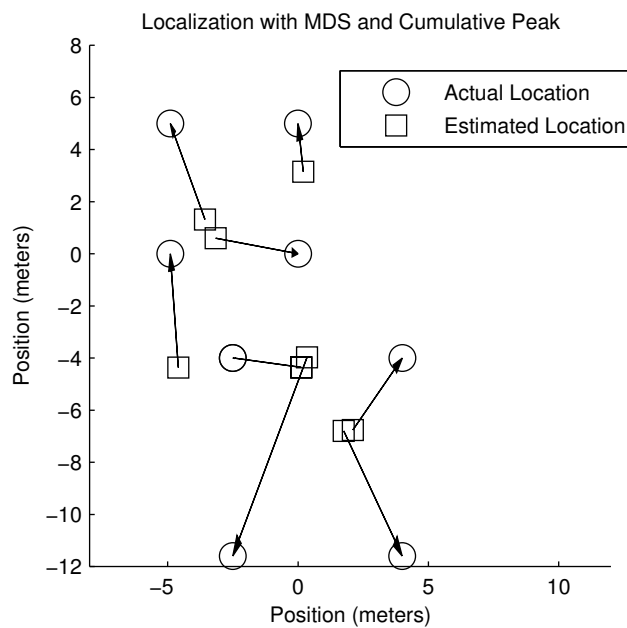

Figure 14 Localization using the peak of the averaged long-time CCF. Localization is worse but not impossible when multipath is not exploited

channels between 14 transmitters and 14 receivers in a simulated cubical office environment with diffuse scattering. In the real measurements, the radio channels between eight transmitters and eight receivers were measured using the RUSK Stanford channel sounder, operating at 2.45 GHz with a bandwidth of 240 MHz. The experimental equipment is special and favors our localization approach. Realistic implementation would require several practical aspects to be considered. Most importantly, a master node would be necessary to centralize the computation and synchronize the receiver nodes. However, using our equipment, we have demonstrated the feasibility of correlation-based radio localization techniques.

Despite the lack of a large number of transmitting antennas, we were able to utilize the spatial diversity of the strongly scattering room by using our improved estimation methods. The main result is that with the real data we were able to estimate spatial antenna locations with less than 2 m error when the theoretical resolution limit is 1.25 m.

Endnotes

[a]These weights are not to be confused with the weighting of the peaks.

Appendix

A.1 Ray-tracing Simulations

Ray-Tracing (RT) is a site-specific geometrical technique that evaluates propagation paths generated by rays as they interact with the environment. A key feature of indoor propagation channels is diffuse scattering. For this reason, in this work, we model the channel with a classic 3-D RT tool [23], improved with penetration and diffuse scattering [24]. The model of diffuse scattering is described in [25]. A geometrical description of the environment, frequency, number of interactions and dielectric properties of materials are some of the input parameters of a RT tool. In the following sections, the ones used in this work are discussed.

A.1.1 Setup Parameters

The simulation frequency has been set to 2.45 GHz. Antenna radiation patterns are the ones of vertically polarized dipoles both at receive (Rx) and source/transmit (Tx) side. Whenever Tx are placed at walls, they radiate only into the relative half-space. A maximum of three reflections, single diffraction, and single-bounce scattering has been used in the simulations. A directive scattering pattern model with scattering coefficient $S = 0.4$ and beamwidth $\alpha_r = 4$ has been chosen. These paths were filtered using a rectangular filter in frequency domain with a bandwidth of 240 MHz to resemble the measurements.

A.1.2 Simulated Environment

A cubicle-style office scenario has been used as input for the RT tool. The dimensions of the room have been set to 50×20 m with a height of 4 m. The walls, the ceiling, and the floor are supposed to be made of concrete. Cubicles are $4 \times 3 \times 1.8$ m and are organized in two rows. Cubicles are represented by their metallic frames that have been treated as perfect electric conductors. A number of 14 receiving nodes were placed in these cubicles, while the ambient noise signals are generated by 14 sources placed at the outer wall of the room. At both sides, the antennas are placed at an height of 1 m. Figure 4 shows a 2-D map of the simulated environment as well as the positions of the receivers and sources. A rough model for the human body as a rectangular parallelepiped has been used. For the human body, a classic two-thirds muscle homogeneous model [26,27] has been used to get realistic values. The time variance of the channel has been modeled by randomly placing ten persons in the scenario in 100 different realizations. The relative dielectric permittivity ε_r was set to 9 for concrete walls, and 35.2 for the human body, while the conductivity σ was set to 0.06 and 1.16, respectively.

A.2 Radio Channel Measurements

In this paper, we use channel measurements obtained during the Stanford July 2008 Radio Channel Measurement Campaign. More details on the full campaign can be found in [14]. In this appendix, we briefly summarize the most important features of the measurement setup.

A.2.1 Environment

To provide good input data for our localization algorithms, we set up the test environment as shown in Figure 10. We took measurements in a cubicle-style office environment with rich scattering due to the metallic frames of the cubicles and highly reflective walls. The room size was around 34×12 m. Eight receivers were placed in two squares, while the transmitters were positioned at the outer walls. To simulate real time-variant environments, people were moving in the room while the measurements were being recorded.

A.2.2 Measurement Equipment

The measurements were taken with the RUSK Stanford channel sounder at a center frequency of 2.45 GHz with a bandwidth of 240 MHz, and a test signal length of 3.2 μs. The transmitter output power of the sounder was 0.5 W. A rubidium reference in the transmit (Tx) and receive (Rx) units ensured accurate timing and clock synchronization. The sounder used fast 1×8 switches at both transmitter and receiver, enabling switched-array MIMO channel measurements of up to 8×8 antennas, i.e., 64 links. The Tx and Rx antennas were off-the-shelf WiFi antennas, which were connected to the switches of the sounder units using long low-loss cables.

The full 8×8 channel was sounded every 100.76 ms. We recorded a total of $T = 1, 200$ samples, capturing the time variations of the channel. By proper calibration,

we removed the RF effects of the equipment and of the cables so that the resulting data only contain the impulse responses of the channels, denoted as $h_{kl}(t, \tau)$.

Author details
[1]Rice University, Houston, Texas, USA [2]FTW Forschungszentrum Telekommunikation Wien, Austria [3]Université catholique de Louvain, Belgium [4]Stanford University, Stanford, USA

Competing interests
The authors declare that they have no competing interests.

References
1. J Chen, K Yao, R Hudson, Source localization and beamforming. IEEE Signal Processing Magazine. **19**(2), 30–39 (2002). doi:10.1109/79.985676
2. N Patwari, J Ash, S Kyperountas, A Hero III, R Moses, N Correal, Locating the nodes: cooperative localization in wireless sensor networks. Signal Processing Magazine, IEEE **22**(4), 54–69 (2005)
3. JA Costa, N Patwari, AO Hero III, Distributed Weighted Multidimensional Scaling for Node Localization in Sensor Networks. ACM Transactions on Sensor Networks **2**, 39–64 (2006). doi:10.1145/1138127.1138129
4. C Hoene, J Willmann, Four-way TOA and software-based trilateration of IEEE 802.11 devices. in *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008*, 1–6 (2008)
5. D Humphrey, M Hedley, Super-Resolution Time of Arrival for Indoor Localization. in *IEEE International Conference on Communications, 2008. ICC '08*, 3286–3290 (2008)
6. Z Low, J Cheong, C Law, W Ng, Y Lee, Pulse detection algorithm for line-of-sight (LOS) UWB ranging applications. Antennas and Wireless Propagation Letters, IEEE **4**, 63–67 (2005). doi:10.1109/LAWP.2005.844145
7. S Galler, W Gerok, J Schroeder, K Kyamakya, T Kaiser, Combined AOA/TOA UWB localization, in *Communications and Information Technologies, 2007. ISCIT '07. International Symposium on*, 1049–1053 (2007)
8. H Lim, LC Kung, JC Hou, H Luo, Zero-Configuration, Robust Indoor Localization: Theory and Experimentation, in *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, 1–12 (2006)
9. M Nezafat, M Kaveh, H Tsuji, Indoor Localization Using a Spatial Channel Signature Database. Antennas and Wireless Propagation Letters, IEEE. **5**, 406–409 (2006)
10. J Garnier, G Papanicolaou, Passive Sensor Imaging Using Cross Correlations of Noisy Signals in a Scattering Medium. SIAM Journal on Imaging Sciences. **2**, 396–437 (2009). doi:10.1137/080723454
11. L Stehly, M Campillo, B Froment, RL Waver, Reconstructing Green's function by correlation of the coda of the correlation (C^3) of ambient seismic noise. Journal of Geophysical Research **113**, 1–10 (2008)
12. WS Torgerson, Multidimensional Scaling: I. Theory and Method. Psychometrika **17**(4), 401–419 (1952). doi:10.1007/BF02288916
13. JC Gower, Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika **53**(3 and 4), 325–338 (1966)
14. N Czink, B Bandemer, G Vazquez-Vilar, L Jalloul, A Paulraj, Stanford July 2008 Radio Channel Measurement Campaign, (Tech. rep., Stanford University, presented at COST 2100, TD(08)620, Lille, France, 2008)
15. J Garnier, G Papanicolaou, Resolution analysis for imaging with noise. Inverse Problems **26**(074001), 1349–1352 (2010)
16. R Parisi, A Cirillo, M Panella, A Uncini, Source localization in reverberant environments by consistent peak selection. in *Proc. of the IEEE ICASSP 2007*, Honolulu, Hawaii (2007)
17. J Spiesberger, Finding the right cross-correlation peak for locating sounds in multipath environments with a fourth-moment function. J Acoust Soc Am. **108**, 1349–1352 (2000). doi:10.1121/1.1287709
18. J de Leeuw, *Recent Developments in Statistics*, (Amsterdam: North Holland Publishing Company, 1977)
19. K Langendoen, N Reijers, Distributed localization in wireless sensor networks: a quantitative comparison. Computer Networks **43**, 499–518 (2003). doi:10.1016/S1389-1286(03)00356-6
20. T Jia, RM Buehrer, A New Cramer-Rao Lower Bound for TOA-based Localization, in *Proceedings of IEEE Military Communications Conference*, 1–5 (2008)
21. M Nicoli, D Fontanella, Fundamental Performance Limits of TOA-based cooperative Localization, in *IEEE International Conference on Communications*, 2009 (2009)
22. S Berger, A Wittneben, Carrier Phase Synchronization of Multiple Distributed Nodes in a Wireless Network. in *8th IEEE Workshop on Signal Processing Advances for Wireless Communications (SPAWC)*, Helsinki, Finland (2007)
23. C Oestges, B Clerckx, L Raynaud, D Vanhoenacker-Janvier, Deterministic Channel Modeling and Performance Simulation of Microcellular Wide-Band Communication Systems. IEEE Trans Veh Technol. **51**(6), 1422–1430 (2002). doi:10.1109/TVT.2002.804846
24. F Mani, C Oestges, Evaluation of Diffuse Scattering Contribution for Delay Spread and Crosspolarization Ratio Prediction in an Indoor Scenario, in *4th European Conference on Antennas and Propagation - EuCAP*, Barcelona, Spain (2010)
25. V DegliEsposti, F Fuschini, EM Vitucci, G Falciasecca, Measurement and Modelling of Scattering from Buildings. IEEE Trans Antennas Propagat. **55**, 143–154 (2007)
26. A Ruddle, Computed SAR levels in vehicle occupants due to on-board transmissions at 900 MHz. in *Antennas Propagation Conference, 2009. LAPC 2009. Loughborough* 137–140 (2009)
27. PS Hall, Y Hao, *Antennas and propagation for body-centric wireless communications* (Artec House, London, 2006)