Radio Propagation-Aware Distance Estimation Based on Neighborhood Comparison

Carsten Buschmann[1], Horst Hellbrück[1], Stefan Fischer[1], Alexander Kröller[2], and Sàndor Fekete[2]

[1] Institute of Telemactics, University of Lübeck, Germany
{buschmann, hellbrueck, fischer}@itm.uni-luebeck.de
[2] Institute of Mathematical Optimization, Braunschweig University of Technology, Germany
{a.kroeller, s.fekete}@tu-bs.de

Abstract. Distance estimation is important for localization and a multitude of other tasks in wireless sensor networks. We propose a new scheme for distance estimation based on the comparison of neighborhood lists. It is inspired by the observation that distant nodes have fewer neighbors in common than close ones. Other than many distance estimation schemes, it relies neither on special hardware nor on unreliable measurements of physical wireless communication properties like RSSI. Additionally the approach benefits from message exchange by other protocols and requires a single additional message exchange for distance estimation. We will show that the approach is universally applicable and works with arbitrary radio hardware. We discuss related work and present the new approach in detail including its mathematical foundations. We demonstrate the performance of our approach by presenting various simulation results.

1 Introduction

In sensor networks self-localization is an important part of self-organization [1]. Location information is essential for the detection of events and the tagging of these events with their geographic origin. Considering a temperature sensing network where some nodes experience a sudden temperature rise, nodes can identify the bounding region of this incident with the help of their position. However, GPS receivers are costly and energy consuming, and thus conflict with major sensor node design goals: low price, tiny form factor and minimal energy consumption. Hence not all nodes can be equipped with GPS, and other techniques have gained popularity. These algorithms assume that a small number of devices, so called anchors, already know their location. The other nodes try to estimate their distances to the anchors and then infer their position through multilateration. Like this, distances became one of the key foundations for location estimation based on anchor nodes [13]. Furthermore, proximity is an intrinsically interesting property of environmental elements and other wireless sensor nodes.

Many different systems for distance estimation (often also called ranging) have been developed. Nearly all of them employ a sender-receiver-scheme: one node

K. Langendoen and T. Voigt (Eds.): EWSN 2007, LNCS 4373, pp. 325–340, 2007.

emits some kind of signal, the other uses a special receiver to measure physical signal properties such as attenuation or time of flight. These techniques will be discussed in detail in the next section. If no direct range distance estimation is possible, multi hop accumulation of distance estimates along the route from anchor to the node is a common work-around (compare e.g. DV-Distance in [15]).

In this paper we present the neighborhood intersection distance estimation scheme (NIDES). This novel approach to distance estimation does not rely on special hardware or unreliable measurements of physical signal properties. Instead, it computes the distances from intersection of sets of adjacent nodes. It is based on the observation that close sensor nodes share more neighbors than distant ones. However, we show that the number of shared neighbors also depends on the radio propagation properties. Hence, we developed a scheme that is radio aware. We will show that the approach is generally applicable and will work with arbitrary radio models.

The remainder of this paper is structured as follows. In the next section we present related work that deals with distance estimation or employs neighborhood lists. We review different distance estimation techniques and discuss their advantages and disadvantages. In Section 3 we give an overview of different radio models and show how their properties can be expressed in so called radio model functions or probability functions. Section 4 then introduces the radio model aware neighborhood intersection distance estimation scheme. We elaborate on its mathematical foundations and discuss properties and factors of influence. In addition, we point out how NIDES can be implemented as a distributed network protocol. Section 5 reports the results of the simulative performance evaluation. The paper is concluded by a summary and directions for future work.

2 Related Work

A number of techniques to acquire distance estimates between nodes have been developed in the past.

One of the first approaches for distance estimation in ubiquitous computing systems was proximity detection. Examples are the Active Badge [21] and the Hybrid indoor navigation system [5]. They both use infrared light to periodically transmit beacons with unique identification. The recipients of these beacons - either the fixed infrastructure or the mobile devices - deduce proximity to the beacon since the range of infrared signals is limited to a few meters of distance. With increasing number of beacons a fine grained resolution is achievable. The downside of this distance estimation technique is the very high deployment effort and the need for a backbone infrastructure.

To enable infrastructure-less distance estimations, a broad range of techniques have been proposed, deployed and researched in the past, which can be classified into the following three categories: approaches based on (differential) time of flight, radio signal strength and connectivity.

The measurement of the *time of flight (ToF)* of a signal is a robust method to estimate distances which is used e.g. by GPS [12]. However, measurements require a tight time synchronization of sender and receiver. Systems like Calamari [22], Cricket [18], AHLoS [20] and others use a technique called "differential time of arrival" (DTOA) to avoid complex time synchronization. They send out two signals (usually ultra sonic sound and radio frequency signals) propagating with different speeds and measure the difference in time of arrival. If both signal propagation speeds are known, a distance can be derived from the delta of arrival. The raw difference measurements tend to yield average estimation errors of about 74% [22]. Yet, quite good accuracies can be achieved by post-processing of the measured data with techniques like noise canceling, digital filtering, peak detection and calibration. While some authors report average range estimation errors of 10% [22], others claim an error of about 1% at a maximum range of 9 m [19]. In [23] the authors report the increase of the maximum range to 12 m with an error of 0.5%.

While these systems yield low estimation errors they have two limitations that confine their applicability in real world deployments. The first is their limited coverage: They are typically able to cover 3–15 m [20] which is only a fraction of the communication range of radio frequency transmitters. The second and far more severe is that they require a separate sender/receiver pair, which implies negative effects on size, cost and energy consumption. All three collide with the aim to design tiny, cheap and highly energy efficient nodes for wireless sensor networks [1].

In order not to require additional hardware, different other schemes have been developed using the radio interface itself to infer distances to other nodes. These approaches use radio signal attenuation properties to model the distance between nodes as a function of the *received signal strength indicator (RSSI).* Systems that rely on the RSSI as input parameter such as [4], [3], [2] tend to be quite accurate for short ranges if extensive post-processing is employed, but are imprecise beyond a few meters [14]. At short ranges, distance estimations exhibiting errors of about 10% at the maximum range of about 20 m [23] are feasible. The uncertainty of the radio propagation imposes problems like multipath propagation, fading and shadowing effects as well as obstacles in the line-of-sight. These effects complicate the development of a consistent model [20]. As a result systems relying exclusively on RSSI values remain inaccurage distance estimators [8].

The idea presented in this paper is based on counting neighbors. Whether communication with nearby node is possible depends also on a RSSI threshold but NIDES inherently compensates RSSI inaccuracies by incorporating links to many neighbors and merging the results instead of measuring the unreliable RSSI values of a single link.

Another improvement of RSSI is presented in [14] where *radio interferometry* techniques are used to achieve an average localization error of 3 cm and a range of up to 160 m with a largest error of approximately 6 cm (which is about 0.04%). However, radio interferometry seems to be susceptible to the effects of shadowing, reflection and multi-path propagation. The downsides of the approach are the high computational complexity of the algorithm and that it requires special features of the radio chip. In addition the accuracy is achieved by long lasting measurements, which renders the approach impractical for mobile scenarios.

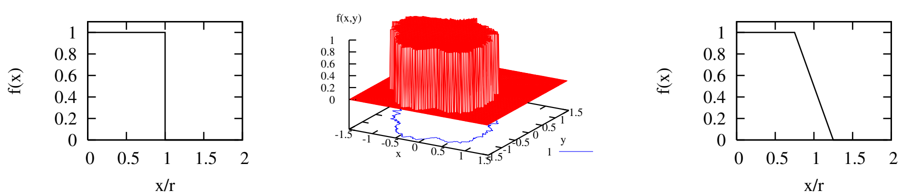
In [6], the authors present the idea to detect far away neighbors by comparing neighborhood lists. However, by determining the percentage of shared neighbors they do not estimate distances. Instead, that fraction influences the probability of forwarding a received flooding message. The fewer neighbors the sender and the receiver share, the higher the forwarding probability is. Thus, distant nodes have a higher probability to forward packets, increasing the coverage per packet sent. However, the authors limit usage of the neighborhood list to forwarding purposes.

3 Radio Model Properties

Not only due to the lack of hardware in real quantities but also for the sake of simplicity researchers in the area of wireless sensor networks are forced to use simulations for validation of algorithms and protocols. Therefore a number of radio models have been developed in the past that express the physical layer characteristics of the radio channel. The key characteristic to determine whether a node is able to receive a radio packet transmitted from a node are the location of sender and receiver, some models also include the status of the radio channel (busy, free with a noise level).

The unit disk graph model (UDG) is the most common model in wireless simulations and available in all network simulators due to its simplicity. It is backed by the observation that (for two dimensional scenarios) the signal strength of wireless communication transceivers fades with the square of the distance from the sender. Given a minimum signal strength required for reception, nodes within the radius of a disk can receive frames from sending nodes. It is rotationally symmetric, and hence yields bidirectional links. Figure 1(a) depicts the radio model function of the unit disk graph model. Here and in all following figures, the communication range has been normalized to 1.

However, experiments have shown that especially in wireless sensor networks with their simple transceivers fading is not circular [9], [25]. To cope with irregularities in the transmission range that is common to real world networks different new models have been proposed. The DOI-Model (DOI means for degree of irregularity) [10] introduces two limits: below the minimum range communication



(a) Unit disk graph model (b) Radio irregularity model (DOI=0.01) (c) Stochastic model

Fig. 1. Probability functions of different radio models

is always possible, above the maximum range communication always fails. In between the two the communication range varies depending on the angle. The DOI factor determines the maximum change in transmission range per degree and thus controls the irregularity of the radio shape (also see Figure 1(b)). In contrast to the circular shape of the unit disk graph model, the DOI model also yields unidirectional links. The radio irregularity model (RIM) [25] further extends the DOI-Model with individual variable sending power that influences the radius of the DOI-shape.

The radio irregularity model is still an oversimplification of the existing radio characteristics because wireless links tend to be error prone so that only a portion of packets is received successfully. Note that for the DOI model nodes within the transmission area receive 100% of the transmitted packets. Measurements in [9] show that packet reception probability decreases with distance but it has a fairly long tail (i.e. even nodes at distances far beyond r may occasionally receive packets). To address this fact Zuniga and Krishnamachari [26] developed a statistical radio model that describes a function for receiving probability fading with the distance. Like the radio irregularity model, it has two bounds. While reception is always possible below the first one, it never takes place above the second. Other than RIM it is rotiationally symmetric but still yield unidirectional links. Figure 1(c) shows a so called linear stochastic model where the change of the packet reception ratio is modelled by a straight line. The bounds are set to $0.75r$ and $1.25r$ here.

There are many other less commonly used radio models, including models a special areas e.g. for cellular network planning. However, these do not provide universal results and are not considered hence.

We will show that NIDES is a general approach for distance estimation that is able to work with arbitrary radio models including unit disk graph, RIM and stochastic radio models.

4 Radio Model Dependent Distance Estimation

The basic idea of the neighborhood intersection distance estimation scheme (NIDES) is to estimate the distance between neighboring nodes from the percentage of neighbors they share. It is based on the observation that two nodes have most of their neighbors in common if they are located close to each other while distant nodes will share fewer or no neighbors at all. For simplification of the model construction we assume a locally uniform node distribution and derive a relation between the percentage of common neighbors of two nodes and their distance. We will address this issue later on in Section 5.

We will show that the fraction of shared neighbors changes with the chosen radio model and will present a generic way to estimate distances with an arbitrary model.

For three different radio models illustrated in Figure 1, an example scenario with the shared neighbors (filled black circles) of two nodes (with an extra black ring) is depicted in Figure 2. Note that the two black nodes feature the same distance in all three cases, variations in the number of common neighbors date back to the radio model. In Figure 2(a) the unit disk graph radio model was employed. The two disks with radius r are indicated as thin circles. The two nodes share 26 neighbors. Figure 2(b) shows the use of the radio irregularity model, the two nodes have 27 common neighbors. The stochastic model was used in the scenario in Figure 2(c), the pair of circles around the two nodes indicates distances of $0.75r$ and $1.25r$. Here, 23 neighbors are shared.

It is evident that the fraction of common nodes varies with the chosen radio model. As a result, distance estimation based on neighborhood intersection without taking the radio model into account will be incorrect. The model should rather be considered as an integral factor of influence for neighborhood-based distance estimation. Hence we propose a scheme that is radio model-aware and incorporates its properties.

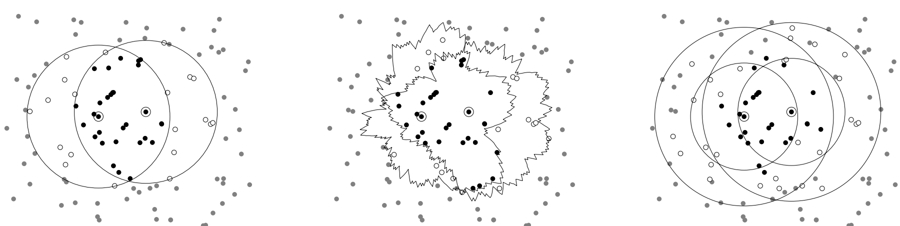
Neighborhood-based distance estimation requires a mapping between the fraction of shared neighbors s of two nodes and their expected distance d from each other that takes the radio model into account. In the following paragraphs we show how a distance estimation function $d(s)$ can be constructed from a given radio model, here described as a two-dimensional function.

The starting point of the construction is the radio model function $f : \mathbb{R} \times \mathbb{R} \to [0, 1]$. It describes the probability that a node that has a position offset of (x, y) to another node can receive its radio signals.

Some radio models provide only a one-dimensional function $f_1 : \mathbb{R} \to [0, 1]$ that models the communication probability as a function of the distance between two nodes (i.e. the radio model is rotationally symmetric). Examples are the unit disk graph or the stochastic radio model. In these cases, a two-dimensional function $f(x, y)$ can easily be constructed by rotating f_1 around the z-axis. The resulting function looks as follows:

$$f(x, y) = f_1(\sqrt{x^2 + y^2}) \tag{1}$$

The created function now also describes the probability that a node located at (x, y) can receive packets from the node at $(0, 0)$.



(a) Unit disk graph model

(b) Radio irregularity model (DOI=0.01)

(c) Stochastic model

Fig. 2. Number of common neighbors varying with the employed radio models

An example for f is shown in Figure 1(b). It depicts the communication probability function of the radio irregularity model with the degree of irregularity set to 0.01.

The estimation function $d(x)$ can now be calculated by convolving f. Because in the general case f is not rotationally symmetric, it must also be rotated during the convolution.

$$f_{rot}(x, y, \alpha) = f(x\ cos(\alpha) + y\ sin(\alpha), -x\ sin(\alpha) + y\ cos(\alpha)) \tag{2}$$

describes f rotated around the z-axis by α. It is translated along the x-axis by d by subtracting d from x. (d will be the distance between the nodes).

The fraction of shared neighbors over distance is now given by $s : \mathbb{R}^+ \to [0, 1]$:

$$s(d) = \int_0^{2\pi} \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{rot}(x - d, y, \alpha)\ f_{rot}(x, y, \alpha)\,\mathrm{d}x\,\mathrm{d}y}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{rot}(x, y, \alpha)\,\mathrm{d}x\,\mathrm{d}y}\,\mathrm{d}\alpha \tag{3}$$

While we here integrate from $-\infty$ to ∞ for the sake of generality, the limits can be adopted to the area of the radio model that adds non-zero values to the integral.

If the radio model is rotationally symmetric like the unit disk graph or the stochastic radio model, the rotation within the convolution can be omitted. The function $s(d)$ is then given by

$$s(d) = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - d, y) f(x, y)\,\mathrm{d}x\,\mathrm{d}y}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}x\,\mathrm{d}y}. \tag{4}$$

In both cases $s(d)$ gives the expected fraction of shared neighbors for a certain distance d.

Figure 3 depicts the result for $s(d)$ of different radio models. For the radio irregularity model, it is depicted in Figure 3(a), the unit disk graph and stochastic model are shown in subfigure (b) and (c).

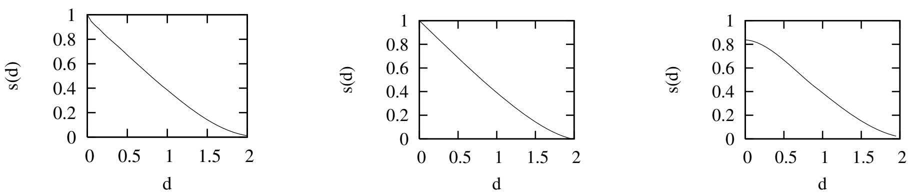
Because s describes the fraction of shared neighbors depending on the distance, its inverse function is required for distance estimation:

$$d(s) = s^{-1}(d) \tag{5}$$

To obtain unique results, s_n needs to be monotonically decreasing. This is the case for all presented radio models.

So far we have shown how the distance estimation function $d(s)$ can be constructed from the radio model probability function $f(x, y)$. In order to employ NIDES in a real world wireless network, basically three steps must be taken:

1. The behavior of the employed communication hardware must be modeled, i.e. the radio model probability function must be determined. This can e.g. be done by conducting measurements. An alternative might be to consult data sheets of the manufacturer.



(a) Radio irregularity model (b) Unit disk graph model (c) Stochastic model

Fig. 3. $s(d)$ for different radio models

2. Once the probability function $f(x, y)$ is known, $s(d)$ can be calculated as described above. It can be used to obtain a lookup table for different distances d, the so called estimation table. It can be stored in the code segment of the target hardware, and hence does not consume any random access memory.
3. At runtime, only neighborhood information must be exchanged between the nodes. When distances need to be estimated, the estimates can be looked up in the estimation table.

Note that the first two steps are done prior to deployment. The generated table is part of the software of the nodes. In this way, every node can estimate the distances to its neighbors if it knows its own neighbors as well as the neighbors of its neighbors. To make sure all nodes have these pieces of information, each node must broadcast two data packets. Due to length limitation of this paper we will only consider distance estimation based on single-hop-neighborhoods here.

First, all nodes must send out a 'hello' packet that enables nodes in range to take notice of them. The sending nodes are inserted into the receivers' neighbor lists. Finally, all nodes have built up their own completed list from the hello packets received. This list must be kept in memory and will be augmented with information about the neighbor distances in the second step. For fast access, the neighbor list should be kept sorted by the node IDs. Unsigned 16 bit values should usually be sufficient to uniquely identify all neighbors. If memory is extremely scarce, neighbor addresses can be easily mapped to at least locally unique (maybe actually 8-bit) identifiers. Even in dense networks with 50 neighbors, this list requires only 100 bytes of memory.

In a second data packet, each node broadcasts its neighbor list. When an adjacent node receives it, it can calculate the number of shared neighbors n_{common}. Because both the received and the local list are sorted, it is enough to iterate once through both lists.

After n_{common} has been acquired, a second value n_{total} representing the overall number of neighbors needs to be determined in order to calculate the fraction s of common neighbors. We propose to employ the average neighbor count, i.e. the mean length of the received and the local neighbor list. Thus fluctuations in the node density and hence in the number of neighbors can be partially compensated for. Note that we do not consider the two nodes to be part of their own neighborhood, and hence omit them when calculating $s = \frac{n_{common}}{n_{total}}$.

Now the receiver can calculate the distance to the sender using $d(s)$ and store the value in a distance list at the index of the sender. Note that it is not required to store the received neighbor list permanently; it can be disposed after the distance has been computed. Only the local neighbor list and the corresponding distances must be kept in memory. To represent the distances, an unsigned 8 bit value should be sufficient.

In some cases, nodes will already have neighbor lists available although NIDES did not explicitly exchange hello messages. Examples for other software that collects neighbor information are routing protocols such as AODV [17], protocols for topology control like SPAN [7] or clustering schemes. When neighbor lists are available from other protocols, the step of broadcasting hello messages can be omitted. Nevertheless, the transmission of the list in the second packet is still required.

5 Simulative Evaluation

To evaluate the distance estimation accuracy of NIDES we ran an extensive set of simulations different scenarios using ns-2 [16]. We implemented the radio irregularity and stochastic communication model by extending the free space radio propagation model.

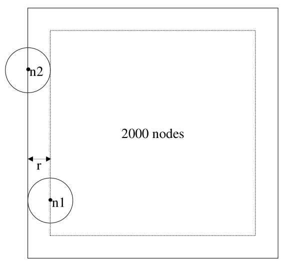


Fig. 4. Simulation scenario

We used the simulation scenario depicted in Figure 4. When analyzing the estimation accuracy, we only considered nodes that were located inside the dotted inner rectangle. The width of this inner area is $2r$ smaller than the simulation area, where r is the communication range. Thus the full communication range of the considered nodes resides within the simulation area (c.f. node n1). This avoids edge effects that would be caused by missing neighbors for node n2 on the left. In order to obtain simulation results for different network densities (i.e. different average neighborhood sizes) we calculated the size of the inner area depending on the desired network density so that a fixed number of nodes resided inside. The size of the full simulation area and the overall node number can then be calculated accordingly. Nodes were spread randomly over the full simulation area, i.e. their x and y coordinates where taken from a uniform random distribution. The average communication range was set to 100.

We iterated over all pairs of adjacent nodes inside the inner rectangle, estimated their distances using NIDES, and compared these estimates to their real Euclidean distances. To obtain statistically sound results, we averaged the results of 100 simulations with the same parameter set.

We focus on simulation results for the radio irregularity model, because it has been developed especially for wireless sensor networks and represents many properties that are typical for this network type. We later present results for the unit disk graph model and the stochastic model.

Depending on radio propagation properties, it is not uncommon that links between nodes are only uni-directional. This effect does not harm the protocol proposed above and also occurs in the simulative evaluation.

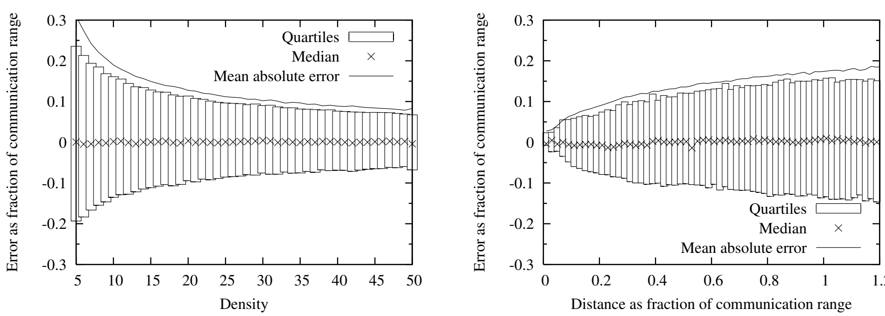
5.1 Radio Irregularity Model

Figure 5(a) shows an interquartile diagram of the estimation error over the average neighborhood size.

In an interquartile diagram, each bar describes a value distribution where only 25% of the values lie above the upper bound and below the lower bound of the bar. Figure 5(a) also includes the mean absolute error and the median. For the earlier the absolute error values were summed up for all considered distances and divided by their number, while for the latter a value is selected so that half of the errors are bigger and the other half is smaller. Hence the median indicates whether the estimates are biased.

The diagram was obtained through simulations with the radio irregularity model and the DOI set to 0.01, the estimation function depicted in Figure 3(a) was used. The error is expressed as a fraction of the communication range r.

The spread of the interquartiles decreases with increasing network densities, i.e. neighborhood sizes. As expected, the estimates get more exact with bigger neighborhoods. With many neighbors, the interquartile of the estimates lies within $0.07r$ around the correct distance (i.e. 50% of the estimates feature an



(a) Error over density (b) Error over distance

Fig. 5. Error characteristic, uniform node distribution, radio irregularity model (DOI=0.01)

error of less than 7% of the communication range). For densities above 10 the mean absolute error is below 20% of the communication range. Note that the median error is always very close to zero which means that NIDES has no tendency at all to systematically over- or underestimate distances.

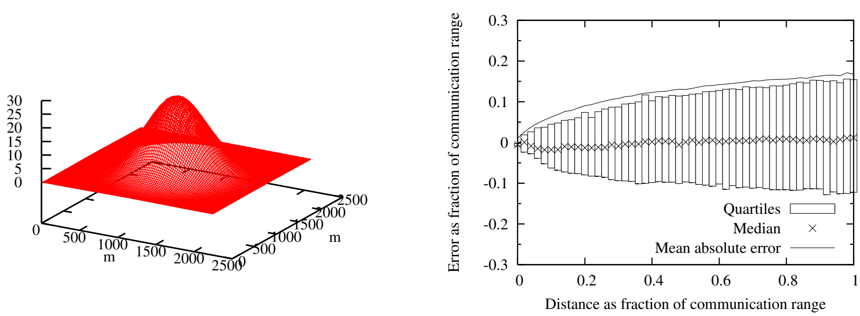
Research has shown that mobile ad-hoc networks tend to partition with high probability if the density, i.e. the average neighborhood size, is below 10 [11]. Others proved that the density required for full connectivity depends on the network size. The more nodes a network consists of, the higher the required density is [24]. For a network comprising 300 nodes, a density of about 13 is needed. Hence, we consider an average neighborhood size of 15 to be a reasonable choice for wireless sensor networks.

With 15 neighbors, the mean error is about $0.15r$ with an interquartile spread of about 0.2. This means that 50% of the estimates feature an error of $0.1r$ or less. If not indicated differently, the simulation results presented in the remainder of the paper feature a density of 15.

We then studied the dependency of the estimation accuracy on the actual distance between the two involved nodes. Figure 5(b) depicts an interquartile diagram of the error over the Euclidean node distance. Both distance and error are expressed as a fraction of the communication range.

The mean absolute error and the interquartile spread increase with the real distance between the two estimating nodes. The reason for that is the decreasing absolute number of shared neighbors. Hence density fluctuation gains importance, and leads to an increased error spread. However, the interquartile spread stays below $\pm 0.15r$ and the mean absolute error is always less than $0.18r$. If these values are averaged over all distances, a mean error of $0.15r$ and an interquartile spread of $\pm 0.1r$ result (as indicated in Figure 5(a) at x=15).

Finally we conducted simulations with non-uniform node distributions. An example of a Gaussian node distribution over the simulation area is depicted in



(a) Gaussian node distribution in the simulation area

(b) Error over distance

Fig. 6. Gaussian node distribution and resulting error characteristic for the radio irregularity model

Figure 6(a). It shows the node density at every point of the simulation area and results from averaging 100 simulations runs. As for the results discussed above, the average node density is set to 15. Different from the scenarios above, we considered all nodes including those at the edges of the network here.

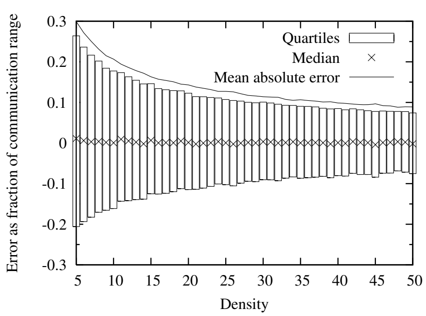
The resulting error characteristic is presented in Figure 6(b). It closely resembles the results obtained with uniform distributions. The error decreases even a bit because of the increased density in the center of the simulation area. The median runs along the x-axis, so obviously the non-uniform distribution does not lead to biased results.

5.2 Linear Stochastic Radio Model

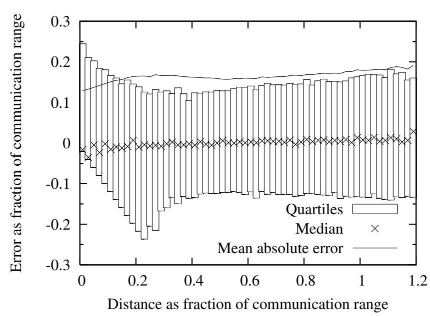
In this subsection we consider the influence of the stochastic model on distance estimation performance. We employed the linear variant with the bounds set to $0.75r$ and $1.25r$ (both as indicated in Figure 1(c)). The corresponding estimation function is depicted in Figure 3(c).

Figure 7 shows the according diagrams. Subfigure 7(a) indicates the distribution of the estimation error depending on the network density (again indicated as a fraction of the communication range r). The mean absolute error is below $0.20r$ for a density of 10 and goes down to 9% of the communication range for a density of 50 while the interquartile spread ranges between $\pm 0.17r$ and $\pm 0.07r$. For a density of 15, the mean absolute error is 16% of the communication range while half of the estimates exhibit an error within $0.12r$ and $0.14r$. The median is always very close to zero.

These values indicate a slightly lower accuracy as if the radio irregularity model was used. The reason for that can be found when looking at Figure 7(b). It indicates the influence of the real distance on the estimation error at a density of 15. While similar to the radio irregularity model for distances above $0.3r$, the stochastic model shows a different behavior below that. The reason is that even if nodes are very close to each other, their neighborhood can vary heavily because the radio model determines a significant fraction of neighbors randomly. This



(a) Error over density



(b) Error over distance

Fig. 7. Error characteristic for the linear stochastic radio model

especially means that even if two nodes are located at the same spot, they will share only 83% of their neighbors on average (compare the distance estimation function depicted in Figure 3(c) at $x = 0$). If nodes have a distance in between each other but share 83% of their neighbors or more, NIDES will estimate their distance as 0, which leads to underestimations for small distances.

In addition, the run of the estimation function is quite flat for small values of x. Thus already minor changes in the fraction of shared neighbors lead to significant changes in the estimated distance. If nodes are close but share slightly less than 83% neighbors of their neighbors, NIDES will estimate a distance relatively far.

Together with the random selection of neighbors, this results in a relatively high error spread for small distances.

However, for small real distances, the negative estimation error is bound by the actual distance of the nodes. Hence also the lower interquartile can be no bigger than the real distance, as visible in Figure 7(b) for distances below $0.23r$.

The increased error for short distances has a negative impact on the overall error. However, this influence is bounded to about 1% of the communication range.

5.3 Unit Disk Graph Model

If the unit disk graph radio model is used, the performance of NIDES is very similar to the case when the radio irregularity model is used.

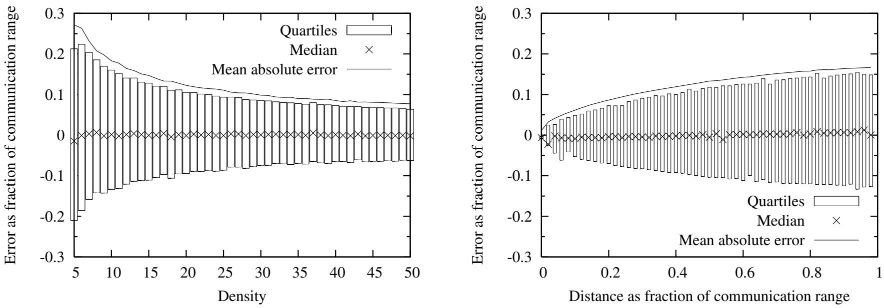
Figure 8 shows the according diagrams. We used the estimation function depicted in Figure 3(b).

Subfigure 8(a) indicates the influence of the average density on the estimation error (again indicated as a fraction of the communication range r). The mean absolute error is $0.27r$ for a density of 5 and goes down to 7% of the communication range for a density of 50 while the interquartile spread ranges between $\pm 0.21r$ and $\pm 0.06r$. For a density of 15, the mean absolute error is 15% of the communication range while 50% of the estimates exhibit an error of below $\pm 0.12r$. These values indicate a slightly higher accuracy as if the radio irregularity model is used. The median is running along the x-axis.

Subfigure 8(b) indicates the influence of the real distance on the estimation error. Again, the behavior of the radio irregularity model is slightly mimicked except for a slightly higher accuracy.

However, the median is not centered for very short distances. It runs on a line that is defined by $g(x) = -x$ instead. This indicates that half of the estimates are 0 even though the nodes have a certain distance, i.e. their displacement is underestimated exactly by their Euclidean distance. The reason is that nodes that are very close to each other have a certain probability of sharing all neighbors. It is due to the fact the fractions of shared neighbors are discrete and only certain values occur (e.g. $\frac{15}{15}$ and $\frac{14}{15}$, but nothing in between) while the real node distance is continuous. This effect occurs again at about $0.55r$ and $0.8r$. Around there, different distances must be mapped on the same fraction.

The effect depends on the node density: the higher the density, the smaller the effect (because the denominator increases). It also appears with the stochastic



(a) Error over density

(b) Error over distance

Fig. 8. Error characteristic for the unit disk graph radio model

and the radio irregularity model but is not as clearly visible because it is averaged out by other stochastic effects. However, it never significantly influences the estimation accuracy.

6 Conclusion and Future Work

In this paper we presented the neighborhood intersection distance estimation scheme (NIDES). This novel approach computes distance estimates from intersection cardinalities of sets of adjacent nodes. In other words, its operation is based on the observation that nodes that are located close to each other share many of their neighbors whereas distant nodes share only few neighbors. While doing so, the presented scheme is radio model aware, i.e. the function that derives the actual distance estimate from the fraction of shared neighbors is constructed from the radio model and hence incorporates the radio properties.

Through simulations we assessed the accuracy of NIDES. While the scheme gets more accurate with increasing network density, it regardless of the radio model yields an average deviation of only about 15% of the communication range when nodes have about 15 neighors.

Thus NIDES classifies in between existing approaches for distance estimation. It achieves a higher accuracy than estimates based on RSSI. Estimates based on differential time of flight feature a higher accuracy than NIDES but demand for specific hardware such as ultra sound emitters and receivers. These have negative effects on cost, size and power consumption, which seems disadvantageous especially for wireless sensor networks. In addition, these have only a very limited maximum range compared to NIDES. Radio interferometry features promisingly low errors but requires specific RF chip properties. Also, this approach needs long series of measurements and complex processing, two disadvantages that can be omitted when neighborhood intersection is used.

We are currently preparing large scale experiments with real sensor network hardware to further validate our simulation results. In the future, we plan to extend NIDES to consider two hop neighborhoods. This enables nodes to sort neighbors with regard to angle, which opens up new vistas for location estimation protocols. In addition, angles along multi hop paths can be estimated. Multi hop distance estimation will also be in the focus of future work.

References

1. I. F. Akyildiz, Y. S. W. Su, and E. Cayirci. Wireless sensor networks: A survey. *Computer Networks*, 38(4):393–422, Mar. 2002.
2. P. Bergamo and G. Mazzini. Localization in sensor networks with fading and mobility. In *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, 2002.
3. J. Beutel. Geolocation in a picoradio environment, 1999.
4. N. Bulusu, J. Heideman, and D. Estrin. Gps-less low cost outdoor localization for very small devices. *IEEE Personal Communications*, 2000.
5. A. Butz, J. Baus, A. Krüger, and M. Lohse. A hybrid indoor navigation system. In *IUI '01: Proceedings of the 6th international conference on Intelligent user interfaces*, pages 25–32, New York, NY, USA, 2001. ACM Press.
6. J. Cartigny and D. Simplot. Border node retransmission based probabilistic broadcast protocols in ad-hoc networks. In *HICSS '03: Proceedings of the 36th Annual Hawaii International Conference on System Sciences (HICSS'03) - Track 9*, page 303, Washington, DC, USA, 2003. IEEE Computer Society.
7. B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. Span: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *Wirel. Netw.*, 8(5):481–494, 2002.
8. E. Elnahrawy, X. Li, and R. P. Martin. The limits of localization using signal strength: A comparative study. In *EEE SECON*, 2004.
9. D. Ganesan, B. Krishnamachari, A. Woo, D. Culler, D. Estrin, and S. Wicker. Complex behavior at scale: An experimental study of low-power wireless sensor networks, 2002.
10. T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher. Range-free localization schemes for large scale sensor networks. In *MobiCom '03: Proceedings of the 9th annual international conference on Mobile computing and networking*, pages 81–95, New York, NY, USA, 2003. ACM Press.
11. H. Hellbrück and S. Fischer. Towards analysis and simulation of ad-hoc networks. In *Proceedings of the 2002 International Conference on Wireless Networks (ICWN02)*, pages 69–75, Las Vegas, Nevada, USA, June 2002. IEEE Computer Society Press.
12. B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System: Theory and Practice*. Springer, 5 edition, 2001.
13. K. Langendoen and N. Reijers. Distributed localization in wireless sensor networks: a quantitative comparison. *Comput. Networks*, 2003.
14. M. Maroti, P. Völgyesi, S. Dora, B. Kusy, A. Nadas, A. Ledeczi, G. Balogh, and K. Molnar. Radio interferometric geolocation. In *SenSys "05: Proceedings of the 3rd international conference on Embedded networked sensor systems*, 2005.
15. D. Niculescu and B. Nath. Ad hoc positioning system (aps). In *Proceedings of GLOBECOM, San Antonio, November 2001.*, 2001.
16. The Network Simulator ns-2 (v2.29). http://www.isi.edu/nsnam/ns/, October 2001.

17. C. Perkins. Ad hoc On-Demand Distance Vector (AODV) Routing. Request for Comments 3561, Network Working Group, Internet Engineering Task Force, July 2003.
18. N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. In *MobiCom '00: Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43, New York, NY, USA, 2000. ACM Press.
19. J. Sallai, G. Balogh, M. Maroti, A. Ledeczi, and B. Kusy. Acoustic ranging in resource-constrained sensor networks. In *International Conference on Wireless Networks*, 2004.
20. A. Savvides, C.-C. Han, and M. B. Strivastava. Dynamic fine-grained localization in ad-hoc networks of sensors. In *MobiCom '01: Proceedings of the 7th annual international conference on Mobile computing and networking*, pages 166–179, New York, NY, USA, 2001. ACM Press.
21. R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. *ACM Transactions on Information Systems*, 10(1):91–102, 1992.
22. K. Whitehouse and D. Culler. Calibration as parameter estimation in sensor networks. In *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*, pages 59–67, New York, NY, USA, 2002. ACM Press.
23. K. Whitehouse, C. Karlof, A. Woo, F. Jiang, and D. Culler. The effects of ranging noise on multihop localization: an empirical study. In *The Fourth International Conference on Information Processing in Sensor Networks (IPSN "05)*, 2005.
24. F. Xue and P. R. Kumar. The number of neighbours needed for connectivity of wireless networks. *IEEE Wireless Networks*, 10(2):169–181, 2004.
25. G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic. Impact of radio irregularity on wireless sensor networks. In *MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 125–138, New York, NY, USA, 2004. ACM Press.
26. M. Zuniga and B. Krishnamachari. Analyzing the transitional region in low power wireless links. In *First IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON)*, October 2004.