Improving Geometric Distance Estimation for Sensor Networks and Unit Disk Graphs

Faisal Aslam, Christian Schindelhauer, and Arne Vater
Department of Computer Science
University of Freiburg, Germany
Email: {aslam, schindel, vater}@informatik.uni-freiburg.de

***Abstract*—Distance measurement between nodes in wireless sensor networks is a prerequisite for a variety of applications and algorithms. However, special hardware allowing such measurements is expensive, especially if dealing with hundreds or thousands of nodes. Fekete et al. presented an approach on distance estimation based on only the neighborhood information available to all nodes in the network. We improve this algorithm, such that it does no longer rely on uniformly distributed nodes. For our approach, it is sufficient that the second derivative of the probability distribution function is a constant.**

I. Introduction

Wireless sensor networks can provide lots of information about the covered area. However, most scenarios (e.g. coverage, routing, tracking, or event detection) require not only the collection of sensors' data, but also the geographic origin of every single measurement, i.e. the position of the sensing node. Unless the nodes are manually placed to exact known locations, the network itself must report the positions of all sensors.

An easy – and expensive – solution is to equip all nodes with special localization devices, e.g. satellite navigation (GPS, Galileo). An alternative is to use distance measurements between neighbored[1] nodes and apply a localization algorithm based on those distances [1]. Note, that although there exists lots of localization schemes that work well in practice, the computational complexity of the corresponding decision problem is NP-hard [2].

The next section gives some examples of the various approaches to measure the distance between two nodes. However, they all rely on special hardware. Other than that, Fekete et al. [3] propose a scheme independent of additional distance measurements. Under the assumption of uniformly distributed nodes, the idea is to count the number of identical neighbors of two different nodes u and v. Based on this and the total amount of neighbors of both nodes, the distance between them can be estimated.

II. Hardware Dependent Localization Algorithms

A distributed localization scheme using noisy range measurements is given in [4]. To overcome the problem of node position ambiguity, the authors consider robust quadrilaterals – subgraphs of four nodes – where all distances between any two nodes are known. Using only such quadrilaterals, it is possible to locate their positions in the original graph with high probability. In case a node's position cannot be localized, it is omitted in the final result, providing a high probability for all other positions to be correct. The prerequisites to use this scheme is a sensor hardware capable of distance measuring (the authors used an additional supersonic sensor together with RF with an accuracy between $1cm$ and $5cm$), and a network density, such that the average degree of nodes is at least 10.

A localization system, which does not require any special hardware capabilities of the nodes except a sensor to recognize light, is called Spotlight and is described in [5]. The idea is an event-driven localization, where nodes recognize an artificially created and well controlled global event. In the described case, this event is a recognizable light (spotlight), whose position and time coordinate is well known. Based on the different times, when the spotlight was recognized by all nodes, the localization is computed. However, an external device for emitting the events is required. This external device has to be aware of its precise position and orientation when the spotlight is emitted making it quite expensive. For example, the device used by the authors of [5] was worth approximately \$1000.

StarDust [6] is yet another localization technique that also uses externally controlled events, a stroboscope light in this case. In contrast to Spotlight, the nodes themselves do not process those signals, but merely reflect them back to the source. The localization is based on a given map of the environment and basically performed by the flashing device. The nodes possibly support the localization with their neighborhood lists.

Both Spotlight and StarDust require a direct line of sight between the event emitting device and the sensor nodes. Another option is to use color filters for the nodes' reflectors to allow a differentiation of single nodes by their color but the accuracy of this method is far less, when compared with Spotlight and StarDust.

III. Model and Overview

In this section we are giving an overview and the basic idea of our new algorithm. A detailed analysis will follow in Section IV.

As communication model, we use the unit disk graph [7], i.e. there exists a link between two nodes, if and only if

[1]Two nodes are neighbors, if they have a direct communication link.

9781-4244-3941-6/09/\$25.00

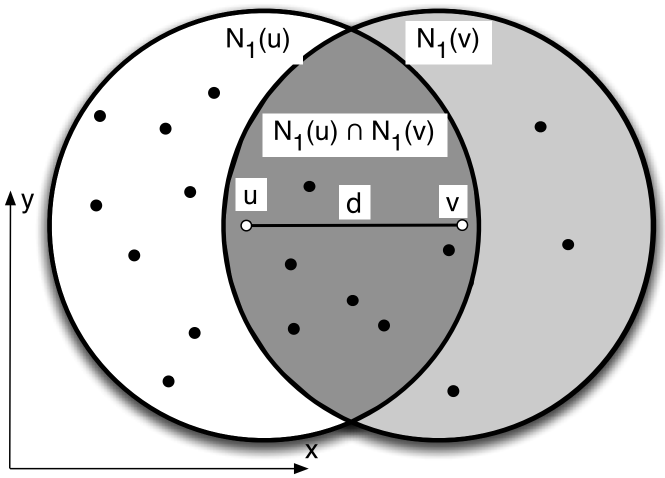

Fig. 1. Distance estimation by counting neighbors.

the Euclidean distance between them is at most 1. Thus, all nodes have the same transmission range. We assume, that all nodes know their two-hop neighborhood by exchanging their neighborhood lists with all adjacent nodes.

Fekete et al. [3] estimate the distance between two nodes by comparing the areas covered by their communication range. The larger the intersection of those two areas, the smaller is the distance between the nodes. Since the size of the areas may be unknown (in the unit disc graph model they depend on the transmission radius), as well as the exact positions are unknown, those areas are estimated by the number of nodes in them, see Figure 1. $N_i(u)$ denotes the i-hop neighborhood of node u. Note, that $N_1(u)$ and $N_1(v)$ are both known to the nodes u and v.

In [3] the authors base the distance estimation between the nodes u and v on dividing one node's covered area by the intersection of both nodes' covered areas: $d = f\left(\frac{|N_1(u)\cap N_1(v)|}{|N_1(u)|}\right)$, where d is the distance.

We argue, that instead of using the intersection $N_1(u) \cap N_1(v)$ in this function as the basis for the estimation of d, it is preferable to use the union $N_1(u) \cup N_1(v)$ instead. Since the considered area is larger, the statistical error is reduced. Furthermore, if one has to cope with inhomogeneous density, the calculated distance between two nodes is symmetrical, which otherwise is not necessarily the case.

IV. Details and Analysis

We continue by giving the details of the distance measurement by our improved algorithm. Afterwards, we compare our approach to the one presented in [3].

A. Distance Estimation Between Nodes

Let $f(x, y)$ denote the two-dimensional probability density function of the positioning scheme in the two-dimensional plane. Furthermore, for simplicity, we identify a node u in the plane by its vector $u = (x, y)^T$ and let $f(u) = f(x(u), y(u))$, where $x(u)$ and $y(u)$ denote the cartesian coordinates of u. All nodes are placed independently.

We assume that the probability density function f is twice differentiable and concentrate on determining the distance $d = ||u, v||_2$ between the nodes u and v. W.l.o.g. we assume that u and v share the same y-coordinate, i.e. $y(u) = y(v)$.

Let $D(u)$ denote the disk with center u and radius 1. Straight-forward geometry proves the following lemma, where $A(R)$ denotes the area of a region R.

Lemma 1: Let $s(d) := 2\arccos\frac{d}{2} - \sin 2\arccos\frac{d}{2}$, and let $|d| < 2$. Then

$$A(D(u) \cap D(v)) = s(||u, v||_2).$$

Thus, the distance between two nodes determines the intersecting area of their communication discs.

The one-sided method in [3] is based on the observation that for all nodes $w \neq u, v$:

$$Pr[w \in D(u) \cap D(v) \mid w \in D(u)] = \frac{\iint_{(x,y)\in D(u)\cap D(v)} f(x, y)dx\ dy}{\iint_{(x,y)\in D(u)} f(x, y)dx\ dy} .$$

If $D(u)$ is within a region where $f(x, y) = c$ is uniform with a constant $c > 0$, it follows

$$Pr[w \in D(u) \cap D(v) \mid w \in D(u)] = \frac{\iint_{(x,y)\in D(u)\cap D(v)} f(x, y)dx\, dy}{\iint_{(x,y)\in D(u)} f(x, y)dx\ dy}$$

This describes the probability of a node being a neighbor of node v if it is already neighbored to node u. The values for $D(u)$ and $D(v)$ are estimated by counting the total number of neighbors. This approach is independent of the communication radius R_c, if R_c is constant for all nodes. However, if the node density in the network (i.e., in the area covered by the two considered nodes) is not constant, one node may have more neighbors than the other, although both nodes cover the same area size, see also Figure 1. In the following, we improve the one-sided estimation method to the symmetric estimation method.

If $D(u)$ is within a region where $f(x, y) = c$ is uniform with a constant $c > 0$, it follows

$$Pr[w \in D(u) \cap D(v) \mid w \in D(u)] = \frac{s(d)}{\pi} .$$

This function is depicted in Figure 2.

If $f(x, y)$ is not constant, the probability can be estimated by the following lemma, if the first derivative $f' = (f'_x, f'_y) := \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$ is a constant:

Lemma 2: Let $f'(u) = (f'_x, f'_y)$ consist of constant functions f'_x and f'_y within the disk $D(u)$. Then, for $y(u) = y(v)$ and $x(v) > x(u)$ it holds:

$$Pr[w \in D(u) \cap D(v) \mid w \in D(u)] = \frac{s(d)}{\pi} \cdot \frac{f\left(\frac{u+v}{2}\right)}{f(u)} = \frac{s(d)}{\pi} \cdot \left(1 + \frac{d \cdot f'_x}{2f(u)}\right)$$

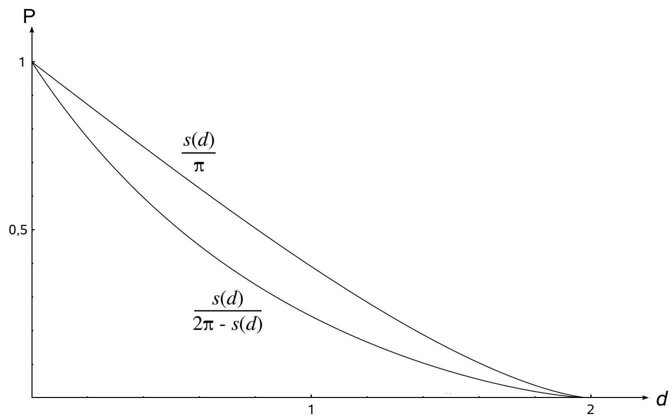

Fig. 2. Illustration of the function $s(d)$.

Proof:

$$f(u+w)+f(u-w)=2f(u)$$

and

$$f(u+w)\in D(u) \iff f(u-w)\in D(v)\ .$$

From this it follows

$$\iint_{(x,y)\in D(u)} f(x,y)dx\ dy = f(u)A(D(u)) = f(u)\pi\ .$$

Using the symmetry point $\frac{u+v}{2}$ for $D(u)\cap D(v)$ we can analogously show

$$\iint_{(x,y)\in D(u)\cap D(v)} f(x,y)dx\ dy = f\left(\tfrac{u+v}{2}\right)s(d)\ .$$

Note that $f\left(\frac{u+v}{2}\right)=f(u)+\frac{d}{2}f'_x$. ∎

So, the one-sided method cannot be used for distance estimation, if the density function is not constant. However, in the special case of a constant derivative one can use the symmetric method.

Lemma 3: If $f'(u)$ is constant within the disk $D(u)$, then

$$Pr[w\in D(u)\cap D(v)\mid w\in D(u)\cup D(v)] = \frac{s(d)}{2\pi - s(d)}\ .$$

Proof: For all $w\in\mathbb{R}^2$ we observe the following:

$$f(\tfrac{u+v}{2}+w)+f(\tfrac{u+v}{2}-w)=2f(\tfrac{u+v}{2})\ ,$$

$$f(\tfrac{u+v}{2}+w)\in D(u)\cup D(v) \iff f(\tfrac{u+v}{2}-w)\in D(u)\cup D(v)$$

and

$$f(\tfrac{u+v}{2}+w)\in D(u)\cap D(v) \iff f(\tfrac{u+v}{2}-w)\in D(u)\cap D(v)\ .$$

From this it follows

$$\iint_{(x,y)\in D(u)\cap D(v)} f(x,y)dx\ dy = s(d)f\left(\tfrac{u+v}{2}\right)$$

and

$$\iint_{(x,y)\in D(u)\cup D(v)} f(x,y)dx\ dy = (2\pi - s(d))f\left(\tfrac{u+v}{2}\right)\ .$$

Dividing these terms proves the claim. ∎

This lemma can be generalized for arbitrary density functions to the following theorem:

Theorem 1:

$$\left| Pr\left[D(u)\cap D(v)\mid D(u)\cup D(v)\right] - \frac{s(d)}{2\pi - s(d)}\right]$$

$$\leq\ \frac{s(d)}{2\pi - s(d)}\cdot\frac{\frac{5}{2}|f''|_{2,\sup}}{1-2|f''|_{2,\sup}}\ ,$$

for $|f''|_{2,\sup} := \sup_{w\in D(u)\cup D(v)}\left|\frac{\partial^2 f(w)}{\partial^2 x},\frac{\partial^2 f(w)}{\partial^2 y}\right|_2$.

Proof: First note that the maximum distance of every point in $D(u)\cup D(v)$ from $\frac{u+v}{2}$ is at most 2. This leads to the following bound for all $w\in\mathbb{R}^2$ with $||w||_2\leq 2$:

$$\left|f(\tfrac{u+v}{2}+w)+f(\tfrac{u+v}{2}-w)-2f(\tfrac{u+v}{2})\right|\ \leq\ |f''|_{2,\sup}|w|^2$$
$$\leq 4|f''|_{2,\sup}\ .$$

Hence,

$$\left|\iint_{(x,y)\in D(u)\cup D(v)} f(x,y)dx\ dy - s(d)f\left(\tfrac{u+v}{2}\right)\right|$$
$$\leq 2(2\pi - s(d))|f''|_{2,\sup}\ .$$

Now for $D(u)\cap D(v)$ the maximal distance to $\frac{u+v}{2}$ is at most 1. This leads to the bound of

$$\left|f(\tfrac{u+v}{2}+w)+f(\tfrac{u+v}{2}-w)-2f(\tfrac{u+v}{2})\right|\ \leq |f''|_{2,\sup}\ .$$

So,

$$\left|\iint_{(x,y)\in D(u)\cap D(v)} f(x,y)dx\ dy - s(d)f\left(\tfrac{u+v}{2}\right)\right|$$
$$\leq \frac{1}{2}s(d)|f''|_{2,\sup}\ .$$

Combining these error bounds gives an error bound for the quotient of $\frac{1+\frac{1}{2}|f''|_{2,\sup}}{1-2|f''|_{2,\sup}}-1$ which proves the claim. ∎

B. Comparison of One-Side and Symmetric Estimators

A natural assumption of a probability density function is a Gaussian distribution with the center $(0,0)$:

$$f(x,y) := \frac{1}{2\pi\sigma^2}\exp\left(-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}\right),$$

where σ is the standard deviation. For this Gaussian distribution we want to compare the quality of the one-sided with the symmetric method. Furthermore we assume that $\sigma>1$.

The error for the one-sided and the symmetric method is maximized for u and v. W.l.o.g. let $y(u)=y(v)=0$. In this case we have the following derivatives:

$$f'_x(x,y) = -\frac{x}{\sigma^2}\ f(x,y)$$
$$f''_x(x,y) = \left(\frac{x^2}{\sigma^4}-\frac{1}{\sigma^2}\right)\ f(x,y)$$
$$f'''_x(x,y) = \left(-\frac{x^3}{\sigma^6}+\frac{2x}{\sigma^4}\right)\ f(x,y)$$

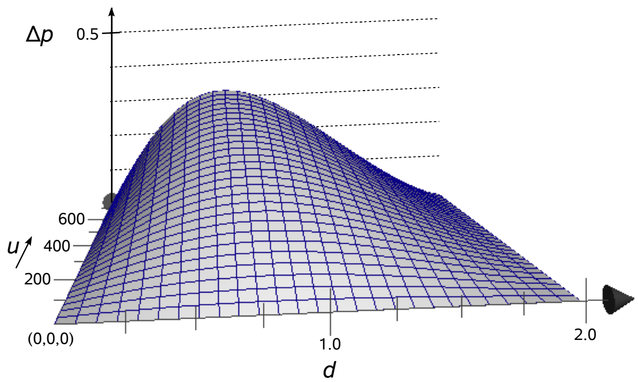

Fig. 3. The graph shows the absolute error difference as probability Δp between the one-sided and the symmetric method for $\sigma = 20$ and $u \leq 750$.

Hence, for $x = \pm\sqrt{2}\sigma$ and $y = 0$, the term $|f''|_{2,\sup}$ is only locally maximal:

$$f''_x(\pm\sqrt{2}\sigma, 0) = \frac{1}{2\pi e \sigma^4}$$

For $x = 0$, $y = 0$ we have the maximum absolute value with

$$f''_x(0,0) = \frac{-1}{2\pi\sigma^4} \ .$$

It follows from Theorem 1:

Corollary 1: For $\sigma \geq 2$ and a (μ, σ)-distributed two-dimensional Gaussian probability density function:

$$\left| Pr\left[D(u) \cap D(v) \mid D(u) \cup D(v)\right] - \frac{s(d)}{2\pi - s(d)} \right|$$

$$\leq \frac{s(d)}{2\pi - s(d)} \frac{3}{2\pi\sigma^4} \ .$$

Proof: Since $\sigma \geq 2$ we have $|f''|_{2,\sup} \leq \frac{1}{32\pi}$. Therefore $\frac{\frac{5}{2}|f''|_{2,\sup}}{1-2|f''|_{2,\sup}} \leq \frac{80}{31}|f''|_{2,\sup} \leq \frac{3}{2\pi\sigma^4}$. ■

If the variance σ^2 of the Gaussian placement distribution is large enough compared to the communication radius $R_c = 1$, then this corollary shows that the symmetric method gives a good approximation of the distance.

Figure 3 shows the deviation from the probability for the one-sided method and the symmetric method, i.e. the absolute difference between the result of the two methods. The graph[2] clearly shows the advantage of the symmetric method over the one-sided method for growing u.

[2]The plotted function in detail:

$$\Delta p = \frac{s(d)}{2\pi - s(d)} \left(\max\left\{ \frac{1 + f''(u+1+d/2)}{1 - f''(u+1+d/2)}, \frac{1 + f''(u-1-d/2)}{1 - f''(u-1-d/2)} \right\} - 1 \right) - \left| \frac{s(d) f(u+d/2)}{(\pi - 2s(d)) f(u) + s(d) f(u-d/2) + s(d) f(u+d/2)} - \frac{s(d)}{\pi} \right|$$

C. Implementation

The computation complexity of the distance estimation by our algorithm is very small. The main calculation that has to be done is the calculation of the function $s(d)$. It merely consists of simple arithmetic functions and, moreover, it can be easily implemented by using a lookup-table to approximate its values. Due to its simple developing, a lookup-table can be kept rather short, cp. Figure 2. Furthermore, based on the distance estimations presented above, a node can also easily estimate its neighboring nodes' relative positions and directions.

D. Sensing Range Based Coverage

Oftentimes localization should be used not solely for communication purpose (e.g. routing), but also to determine the sensing coverage of a network. Since a sensor node's communication range will usually be different from its sensing range, it is not sufficient for all nodes to know their two-hop communication neighbors. In fact, to enable a similar approach for sensing coverage as for communication, the two-hop *sensing neighborhood* must be known. Keep in mind, that both types of coverage should be ensured by local decisions of nodes.

This information is obviously already available, if the communication radius R_c is larger than the sensing radius R_s. However, if $R_c < R_s$, then additional information is required. A straight-forward approach would be to enlarge the collected i-hop neighborhood information $N_i(u)$ of node u, such that $R_s/R_c \leq i/2$. This holds under the assumption that all i-hop neighbors have exactly a geometric distance between $(i-1) \cdot R_c$ and $i \cdot R_c$. This assumption is only true for infinitely dense networks. In real-life scenarios, the average radii of the discs containing the i-hop neighbors around a node become smaller, the sparser the network is.

A simple approach to cope with this problem could be to multiply R_s by a sufficiently large factor to compensate the deviation. The tradeoff would be possibly unnecessarily gathered neighborhood information, i.e. traffic overhead.

V. Conclusion

We took up the idea of hardware independent node localization by Fekete et al. [3] and refined it. We improved the accuracy of this method for scenarios with non-uniform node placement, if the second deviation of the probability density function is constant. Our scheme can especially cope with scenarios, in which the node placement follows a Gaussian distribution, e.g. if all nodes are dropped simultaneously from a plane. Still, distributed implementation is easy and only small computational power is required from each node. Since a higher density (i.e. a higher degree) increases the accuracy of the relation between the amount of common neighbors and the size of the intersecting communication areas, the design is especially suitable for scenarios, in which the communication range of nodes is large compared to the sensing range, if full sensing coverage is desired.

References

[1] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: a quantitative comparison," *Comput. Networks*, vol. 43, no. 4, pp. 499–518, 2003.

[2] J. Aspnes, D. Goldenberg, and Y. R. Yang, "On the computational complexity of sensor network localization," in *Algorithmic Aspects of Wireless Sensor Networks: First International Workshop, ALGOSENSORS 2004, Turku, Finland, July 16, 2004. Proceedings.*, ser. Lecture Notes in Computer Science, vol. 3121. Springer-Verlag, 2004, pp. 32–44.

[3] S. P. Fekete, A. Kröller, C. Buschmann, and S. Fischer, "Geometric distance estimation for sensor networks and unit disk graphs," in *16th International Fall Workshop on Computational Geometry*, 2006.

[4] D. Moore, J. Leonard, D. Rus, and S. Teller, "Robust Distributed Network Localization with Noisy Range Measurements," ACM Conference on Embedded Networked Sensor Systems (SenSys), 2004.

[5] R. Stoleru, T. He, J. A. Stankovic, and D. P. Luebke, "A High-accuracy, Low-cost Localization System for Wireless Sensor Networks," ACM Conference on Embedded Networked Sensor Systems (SenSys), 2005.

[6] R. Stoleru, P. Vicaire, T. He, and J. A. Stankovic, "StarDust: A Flexible Architecture for Passive Localization in Wireless Sensor Networks," ACM Conference on Embedded Networked Sensor Systems (SenSys), 2006.

[7] B. N. Clark, C. J. Colbourn, and D. S. Johnson, "Unit disk graphs," *Discrete Math.*, vol. 86, no. 1-3, pp. 165–177, 1990.