Ad Hoc & Sensor Wireless Networks Vol. 3, pp. 141–170
Reprints available directly from the publisher
Photocopying permitted by license only

©2007 Old City Publishing, Inc.
Published by license under the OCP Science imprint,
a member of the Old City Publishing Group

3D Wireless Network Localization from Inconsistent Distance Observations

RAINER MAUTZ*, WASHINGTON OCHIENG**, GARY BRODIN†, AND ANDY KEMP†

*Geodetic Metrology and Engineering Geodesy, Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, 8093 Zurich, Switzerland
E-mail: mautz@geod.baug.ethz.ch
**Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, United Kingdom
E-mail: w.ochieng@imperial.ac.uk
†Department of Electronic and Electrical Engineering, The University of Leeds, Leeds LS2 9JT
E-mail: g.brodin@leeds.ac.uk, a.h.kemp@leeds.ac.uk

Received: December 1, 2005. In Final Form: April 3, 2006.

This paper presents a 3-dimensional algorithm for determining the node locations in ad-hoc wireless local area sensor networks capable of inter-nodal distance measurement (ranging). In accordance with common classification of localisation strategies, the paper proposes a geometric centralised optionally anchor free algorithm that is based on clusterisation, multilateration and geodetic network adjustment. The method avoids fold-ambiguities by statistical tests on the robustness of the minimally stable structure in 3D. Such a rigid cluster is expanded by a novel robust approach for multilateration. The algorithm developed is capable of obtaining the least-squares solution reliably even in the presence of measurement noise levels of up to 7% of the ranges. Simulations with randomly deployed networks show that it is possible to reach the best achievable mean position deviation (i.e. the Cramer Rao bound).

Keywords: Ad-hoc networks, network localisation, error propagation, network adjustment

1 INTRODUCTION

Precise knowledge of the node locations is an essential component for a diverse set of applications employing ad-hoc wireless local area sensor networks in which nodes have the ability to measure the distance to nearby nodes. Ad-hoc

networks consist of numerous nodes which are deployed arbitrarily (i.e. not in a planed and geometrically optimised fashion) and can grow organically. Unlike in geodetic networks, in some instances the node locations are not known *a priori* – not even approximately before the network is initialised. This paper focuses on the difficult task of initialising a network in the presence of measurement noise. The developed method is based on the quint, a clique of 5 nodes, which is the minimally rigid structure in 3D, and provides statistical tests to filter quints for robustness to prevent large localisation errors.

The choice of the localisation strategy is driven by the user requirements and the system parameters of iPLOT (intelligent Pervasive LOcation Tracking), a Bluetooth based wireless sensor network that is currently being built by Imperial College London in collaboration with the University of Leeds, the UK Home Office Scientific Development Branch, the Forensic Science Services and New Forest Communications Limited. The time of arrival method is used to determine the distance between equipped radio devices. Initial results show that a ranging accuracy of 1 m (1 sigma) is potentially achievable with a projected maximum range of 10 m which can be extended to 100 m by increasing the power. More details about ranging are given in Mautz *et al.* (2006). In summary potential iPLOT users require automatic, accurate, reliable 3D positioning with high integrity and availability for irregular ad-hoc deployed networks with, or optionally without the presence of absolute position reference (Mautz *et al.*, 2006). This paper details a localisation algorithm that attempts to cope with the challenges resulting from the user requirements.

Recently, various localisation systems have been proposed (Goldenberg *et al.*, 2005). While some authors focus on network localisation that can be achieved in theory based on a graph that represents node connectivity, others present their results from actual two dimensional radio networks.

Savvides *et al.* (2003), Savarese *et al.* (2002) as well as Niculescu and Nath (2001) discuss anchor based multi-hop localisation methods, which usually start with a rough estimate of the distances between unknown nodes and some known anchor nodes, then evaluate approximate positions and conclude with an iterative refinement of the positions. This strategy can be classified as topological, because it makes use of node connectivity via multiple-hops making it sensitive to the irregularity of network topologies and the deployment of anchor nodes. Moreover, neither algorithm considers how measurement noise can accumulate and cause incorrect realizations of mirroring ambiguities. Shang *et al.* (2004) have developed an improved scheme for multidimensional scaling (MDS), a localisation method that transforms proximity information into geometric embedding. Since MDS also relies on rough multi-hop distance estimations, its performance is poor for irregular networks and the solutions are subject to local minima (Zhang, 2004).

Only a few approaches take account of errors in range measurements in their localisation strategies. Savvides *et al.* (2003) develop a lower bound on the accuracy for achievable location estimates taking into account different parameters such as ranging errors, node connectivity or anchor density.

Whitehouse *et al*. (2005) study how Gaussian noise based models cope with empirical ranging characteristics. Moore *et al*. (2004) present a robust 2D localisation strategy that is closely related to the method proposed in this paper. The improvements over Moore's robust quads approach are robust ambiguity assessment and enhanced multilateration both resulting in higher integrity. Furthermore, less inter-node connectivity is required and the extension into 3D is provided.

2 KEY ELEMENTS OF THE APPROACH

Existing methods tend to implement 2D localisation algorithms and assume similar results for higher dimensions. The approach developed in this paper is configurable and can handle two and three dimensional scenarios. Our strategy requires central computation at a location server and is therefore not set up for distributed processing. However, it is self-organising and capable to locate nodes within ad-hoc networks. Anchor nodes (also known as beacons or control points) are not required, because the coordinates are optionally transformed into the reference system within the concluding step of the algorithm. There is no assumption made on the maximum distance at which range measurements are possible. Consequently, the algorithm does not use the *Noisy Disk* model, in which the distance estimates are defined to be normally distributed around their true values, but undefined if the distance is larger than a maximum range. The approach taken is *geometric* because it carefully uses the direct (single-hop) distances and exhibits higher traceability than *topological* strategies that start off with rough multi-hop distance estimates to guess initial node positions which are refined later. Topological methods require anchor nodes, node collaboration, evaluation at large scale and don't work very well in irregular networks. The main reason for not choosing a topological method is that the very rough initial position estimates are likely to represent local sub-optimums which typically represent falsely folded parts of the network that cannot be overcome by local refinement. The strategy of the new method is to cautiously deploy the nodes with verified lateration and statistical tests, thereby avoiding local optimums. Furthermore, the proposed algorithm can be categorized as *fine-grained*, as it uses direct ranges compared to *coarse-grained* methods that depend on proximity information.

The main characteristic of our localization strategy is that every single step of the position determination is done in an over-determined or redundant manner. If a step can be carried out consistently but without any redundancy, it is aborted. The power of this approach becomes clear when the existence of errors in real-world measurements is considered. The use of consistent data is tempting, as it is straight forward, computationally inexpensive and unambiguous in the results. However, in an error affected environment a lack of redundancy leads to a lack of control of the errors in the system. As network localisation is often a sequential process, initially small observation errors tend

to accumulate. Additional observations cannot stop the propagation of errors, but introduce inconsistencies in the measurements that allow for integrity and error control. The figures for the system's inconsistencies are the basis for the decisions that are made as the nodes are located, i.e. range observations between nodes can be rejected or the clusterisation of nodes can be prevented.

The localisation strategy can be broken down into three phases:

(I) Creation of clusters
Starting from a graph G that represents network connectivity, five nodes are identified, that are all linked to each other by range measurements. If such an initial cluster passes statistical tests, it is released for further expansion. Additional nodes are added consecutively using a verified multilateration technique.

(II) Merging of clusters
Experiments show that only a fraction of nodes can become a member of one single cluster. The remaining nodes are likely to make up their own clusters which may or may not be connected to neighbouring clusters. In case two clusters share an adequate number of nodes and/or range observations between them, they can be merged using an over-determined 3 dimensional 6-parameter transformation.

(III) Transformation of the cluster(s) into a reference coordinate system
The output of step (II) may again consist of one or more clusters with their individual datum definitions. However, any cluster containing three⋆ nodes with an additional set of reference coordinates is eligible for a transformation into that particular coordinate system.

In the application process, the order of steps is driven by the quality of the data. For instance, if 5 nodes with 'robust' ranges to each other cannot be found, step (I) cannot be performed. In this case the localisation algorithm will follow a different path which we call 'coarse positioning' (Section 3.4). The omission of steps (II) and (III) is also driven by the data. In a well connected obstacle-free network the outcome of step (I) may be one single cluster. Consequently the algorithm can continue directly with step (III) and carry out the transformation. Step (III) however, can only be carried out if sufficient anchor points are available. Omission of step (III) will result in a set of local coordinates – which may have their relevance for several applications.

The optimisation – also known as refinement, relaxation, network adjustment or local gradient method – is an additional tool that may or may not be executed before, between or after any step of the algorithm and is therefore not itemised. The decision to include optimisation is driven by constraints on computational costs and accuracy. However, the generic goal for each step is

⋆Strictly speaking, the case of 3 control points has an intrinsic ambiguity problem which can only be solved with addtional information, e.g. a range observation between the clusters.

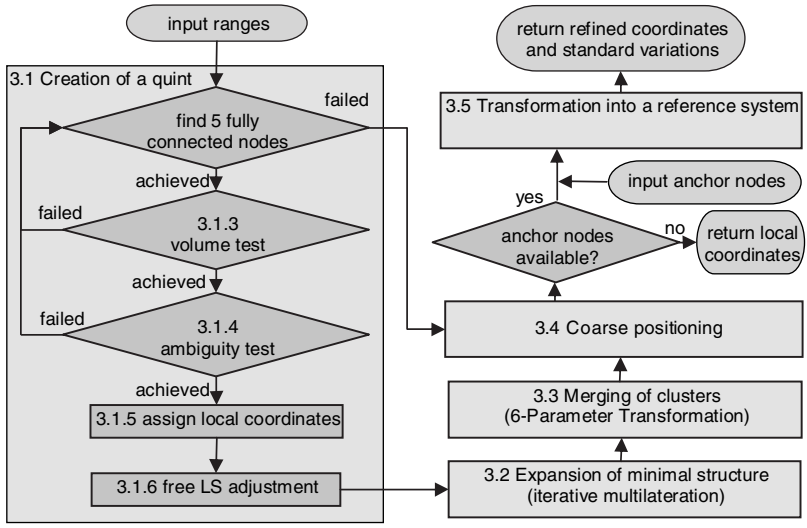

FIGURE 1
Summary of the algorithm phases and their section numbers.

the provision of good quality approximate values for the unknown locations that are necessary for a successful network adjustment.

The localisation approach presented in this paper is centralised. However, in order to reduce the overall complexity in large networks (e.g. more than 100 nodes), the handling of large matrices can be avoided by setting an upper limit for the cluster size.

The main phases of the localisation algorithm and the corresponding section numbers are shown in Figure 1 as a flowchart.

3 ELABORATION

This section details our localisation strategy with a special focus on the first part – the creation of a rigid cluster – which aims to compute the positions of nodes in a local coordinate system that is amenable to global translation, rotation and reflection.

3.1 Creation of the Smallest Redundant Structure

In 3 dimensions, the smallest redundant and rigid graph consists of 5 network nodes with distance constraints to each other. Our approach uses such a structure as a seed for the creation of a larger network. We call this structure quintilateral (or quint). However, before assigning local coordinates to the five participating nodes, we analyse the structure for its robustness and uniqueness with respect to uncertainties in the range measurements.

3.1.1 Ambiguity Problem of a "Quintilateral"

The quintilateral is of relevance for clusterisation because it is the smallest possible subgraph in three dimensions that has only one unique embedding besides global rotations, translations and a reflection. In two dimensions, Moore *et al.* (2004) have approached the cluster localization problem based on a fully-connected point set of 4 nodes, accordingly referred to as "Quadrilateral".

The quintilateral is a fully connected set of 5 nodes as shown in Figure2a. Accordingly, each node is linked to the other 4 nodes by distance measurements, which means the graph of a quint is 4-connected. Since a graph of n nodes has $3n$ degrees of freedom in 3 dimensions, the quint has 15 degrees of freedom. 6 degrees of freedom correspond to 6 datum constraints (3 rotations and 3 translations) making up the total number of degrees of freedom as $3n - 6 = 9$. With the quint holding 10 range measurements, each of them eliminating 1 degree, only 1 redundant or independent observation remains.

The 10^{th} observation is redundant and used to solve for a flip ambiguity. Assume the range r_{45} between node 4 and 5 has not been observed as shown in Figure 2b. In this case, the structure is not rigid as there are two embeddings based on the remaining 9 ranges. As shown in Figure 2c, the node 5 can be mirrored at the base plane containing the nodes 1, 2 and 3 without violating the distance constraints. In other words, the range r_{45} disambiguates between the two possible embeddings (2b) and (2c) with mirror symmetric locations for node 5. For the case as shown in Figure 2, it is obvious that the distances between nodes 4 and 5 differ significantly between (2b) and (2c). But this is not always the case. For some geometric constellations the ambiguity cannot be solved by the redundant range r_{45}, because the difference between the distances d_{45} and $d_{45'}$ is of the same order of magnitude as the ranging error. Such a configuration is shown in Figure 3, where both nodes 5 and 5′ have

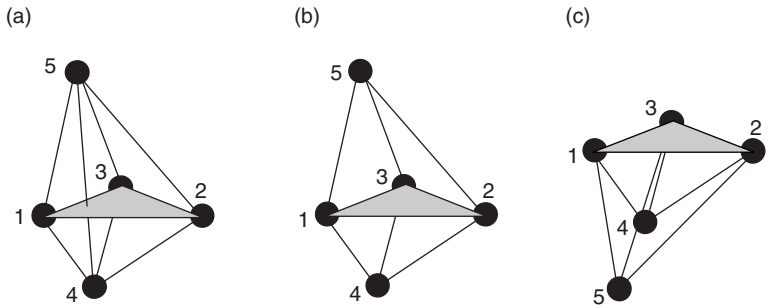

FIGURE 2
(a) Quintilateral, (b) two tetrahedrons T_{1234}, T_{1235}, (c) Tetrahedrons T_{1234} and T'_{1235}, the latter is the reflected version of the tetrahedron T_{1235}.

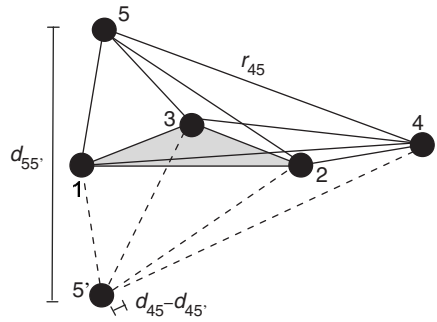

FIGURE 3
Quintilateral with two possible embeddings Q_{12345} and $Q_{12345'}$.

been trilaterated from the nodes 1, 2 and 3. In order to decide which of the two embeddings is correct, the computed distances d_{45} and $d_{45'}$ are compared to the measured distance r_{45}. In Figure 3 however, the difference $|d_{45} - d_{45'}|$ is comparatively small. Consequently, the differences between the measured and the calculated distances $\Delta_{45} = |r_{45} - d_{45}|$ and $\Delta_{45'} = |r_{45} - d_{45'}|$ are both small. A clear decision can only be made, if one of the differences undoubtedly results from measurement noise while the other is significantly above the noise level. Assuming a mean error of the range measurement r_{45}, say 10% of the range, both differences Δ_{45} and $\Delta_{45'}$ in Figure 3 are likely to pass a statistical test on their null hypotheses, which means that both could be a result of noise. Consequently, the range r_{45} does not disambiguate between both embeddings.

Obviously, the problem has its symmetry: if node 5 is defined to be on one particular side of the base plane, the range r_{45} disambiguates between node 4 and its mirrored version $4'$, which has been reflected at the base plane. Note that each quint has its intrinsic mirrored version of itself and therefore it is impossible to decide which of these two mirrored versions is correct. The correct reflection can only be identified if at least 4 points have coordinate values in a reference coordinate system. However, within the given set of 10 ranges only, the decision whether node 4 and 5 are on the same side of the base plane or on opposite sides has to be made. Obviously, critical cases appear when at least one of the nodes 4 or 5 is close to the base plane as exemplified in Figure 3.

Figure 3 shows that finding a unique embedding of a quint can be an ill-posed problem. The decision, whether node 5 or $5'$ is accepted, is based on a small value $|d_{45} - d_{45'}|$. A small error in one of the range measurements can lead to a wrong ambiguity solution and result in a large displacement in the location of node 5, shown in Figure 3 as $d_{55'}$. The best way to deal with this problem is to reject unstable point formations. It is better to reject

a non-robust quint than rely on a structure which may have incorrect internal flips. It is crucial to ensure a correct embedding for several reasons. Firstly, as mentioned before, the displacement of single nodes can be large. Secondly, these errors affect the expansion of the structure severely when additional nodes are added to a falsely flipped quintilateral by multilateration. Thirdly, and most importantly, incorrect flips in a network are difficult to be eliminated later by using optimisation. Refinement algorithms are usually based on iterative local optimisation techniques. Local optimisation however follows the gradient of a multidimensional objective function – which means that it is not possible to get out of a sub-optimum. Unfortunately, incorrect flips reflect such local optimums where local optimisation algorithms converge to. In order to overcome false flips, globally operating optimisation techniques would have to be used, which are computationally expensive and therefore prohibitive for ad-hoc sensor network positioning.

For the reasons given above, we consider the robustness and uniqueness of a configuration – even under the presence of errors in the range measurements – as important. In agreement with Goldenberg *et al.* (2005) we realise that degenerate configurations such as linear alignment of points have zero probability of appearing and the assumption that the network nodes are in general position is justified. However, configurations being close to degenerate may very well occur and therefore affect the network localisation.

Moore *et al.* (2004) have considered the fold ambiguity problem in their 2D localisation algorithm. They analyse a left-right symmetric constellation and estimate the worst-case probability of error in their planar quadrilaterals. Besides, their error propagation relies on derived variables such as angles and not exclusively on the original range measurements.

Our goal however is to estimate the probability of an incorrect flip to occur. Therefore, we develop proper error propagation for the general quintilateral, which is exclusively based on the range measurements. We also do not follow a stepwise error propagation, which would only approximate the propagated observation uncertainties. Approximate formulas can be developed easily, if computation time is critical. In the next section we first justify the use of local coordinates and than continue to detail the ambiguity tests in Sections 3.1.3 and 3.1.4.

3.1.2 Reasons for the Analysis of a "Quintilateral" Without Use of a Reference Frame

Geodetic networks are typically analysed with models involving coordinates, which refer to more-or-less arbitrary reference frames or geodetic datums. As a consequence, the coordinate estimates depend not only on the network shape, but also on its arbitrary position introduced by the choice of a particular reference frame. For this reason it is essential to distinguish between coordinate functions which are estimable quantities and those which are not. As shown by Grafarend and Schaffrin (1976), estimable quantities are those which remain

invariant under changes of the reference frame. Consequently, coordinates depending on a reference frame – including local Cartesian coordinates – are non-estimable quantities.

Since quintilaterals are analysed only by shape, independently of position, rotation and reflection, it is only reasonable that they are studied through properly defined parameters which are independent of frame definitions. Therefore, we abstain from the use of coordinates and restrict the calculations to the original range measurements and barycentric coordinates which are independent of a reference frame.

The computation of a quintilateral is performed in two steps. First, the robustness of the structure is tested under the assumption that the range measurements are corrupted by zero-mean uncorrelated random noise with a Guassian probability density function. These tests aim at the prevention of incorrect realizations caused by flip ambiguities due to the measurement noise. The tests are carried out using barycentric or 'inner' coordinates. Here, nodes 1, 2 and 3 are arbitrarily chosen to define a base plane, where the two general tetrahedrons T_{1234} and T_{1235} that assemble a quint (as shown in Figure 2b) are defined based on frame-independent estimable quantities such as volumes, areas, edge lengths or heights. Only after a quint has passed the tests on robustness, is the second step performed by assigning local coordinates to the nodes. This step realises a transition from datum independent barycentric coordinates to datum dependent local Cartesian coordinates. Next, we continue to describe the robustness tests of a quint.

3.1.3 Volume Test of a General Tetrahedron

A general tetrahedron is considered as robust in the presence of range errors if the uniqueness of its topology is guaranteed up to a certain probability. Clearly, a tetrahedron defined by its 6 edges has two possible configurations corresponding to a reflection across the set of 3 nodes defining the base plane. Based on these 6 edges alone, the correct reflection cannot be identified and an initialisation must begin by arbitrarily choosing one of them. But as soon as the tetrahedron is expanded – as is the case for a quint – the correct reflection needs to be chosen. This choice can only be made properly if the probability of making an incorrect choice can be ruled out up to a certain probability. The determination of this probability is discussed in the next paragraph.

The probability of the occurrence of an incorrect flip in a tetrahedron is given by the probability that one node, say node 4 in Figure 4, is actually on the other side of the base plane in spite of the assumption that has been made. Consequently, critical configurations involve tetrahedrons where the error of the tetrahedron height allows for node 4 to be on both sides with a certain probability – where one instance would reflect the probability of obtaining a negative height. In order to incorporate all four heights and the four triangles

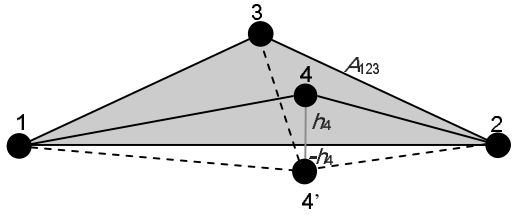

FIGURE 4
Tetrahedron T_{1234} and its reflection$T_{1234'}$ with a negative volume.

in the assessment, we simply determine the probability of having a negative volume of the tetrahedron.

The content C of a simplex in n dimensions is given by the determinant

$$C_n^2 = \frac{(-1)^{n+1}}{2^n (n!)^2} \det \mathbf{M}_n$$

$$= \frac{(-1)^{n+1}}{2^n (n!)^2} \det \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & r_{1,1}^2 & r_{1,2}^2 & \cdots & r_{1,n+1}^2 \\ 1 & r_{2,1}^2 & r_{2,2}^2 & \cdots & r_{2,n+1}^2 \\ \vdots & \vdots & \vdots & \ddots & r_{3,n+1}^2 \\ 1 & r_{n+1,1}^2 & r_{n+1,2}^2 & \cdots & r_{n+1,n+1}^2 \end{pmatrix}, \tag{1}$$

where $\det \mathbf{M}_n = D(1, 2, \ldots, n+1)$ is a $(n+2) \times (n+2)$ matrix known as the Cayley–Menger determinant. $\mathbf{M}_n$ can be obtained from the $(n+1) \times (n+1)$ matrix $\mathbf{M}_n^* = (d_{ij}) = (r_{ij}^2)$ where the elements of $\mathbf{M}_n^*$ represent the squared distances between the nodes, rendered by a top row $(0, 1, 1, \ldots, 1)$ and a left column $(0, 1, 1, \ldots, 1)^{\mathrm{T}}$. For $n = 3$, the volume of the general tetrahedron is given by the determinant

$$V_{n=3}^2 = \frac{1}{288} D(1, 2, 3, 4) = \frac{1}{288} \det \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & r_{12}^2 & r_{13}^2 & r_{14}^2 \\ 1 & r_{12}^2 & 0 & r_{23}^2 & r_{24}^2 \\ 1 & r_{13}^2 & r_{23}^2 & 0 & r_{34}^2 \\ 1 & r_{14}^2 & r_{24}^2 & r_{34}^2 & 0 \end{pmatrix}. \tag{2}$$

The Cayley–Menger determinants have the advantageous property of having a geometric meaning and being mathematically tractable. For example the determinant $D(1, 2, 3, 4)$ permits realising that if $D(1, 2, 3, 4) = 0$, all nodes are deployed in one particular plane. In this case the tetrahedron would be degenerated into a planar quadrilateral. Although such a configuration is unique, it cannot be used to initialise a network and therefore needs to be rejected.

Recall that the presence of small uncertainties in the ranges will cause large uncertainties in the structure.

In the presence of measurement noise, the case $D(1, 2, 3, 4) < 0$ may also occur. As a consequence, the volume of the tetrahedron would be an imaginary number, using (2). Clearly, such a configuration is to be rejected right away since the nodes cannot be located properly. From a geometric point of view, the values of range measurements don't allow for intersection, i.e. when one of the 4 triangles fails its triangle inequality[†].

The error of the volume is determined from range measurements using error propagation. If $\mathbf{r} = (r_1, r_2, \ldots, r_n)$ is a set of random variables with the covariance matrix $\mathbf{C}_r$, and if $\mathbf{F} = (F(r_1), F(r_2), \ldots, F(r_m))$ is a set of transformation functions which are linear or well approximated by the linear terms of a Taylor series in the neighbourhood of the mean $\mathbf{E}(\mathbf{r})$, then the covariance matrix $\mathbf{C}_\mathbf{F}$ of $\mathbf{F}$ is $\mathbf{C}_\mathbf{F} = \mathbf{F} * \mathbf{C}_\mathbf{r} * \mathbf{F}^\mathrm{T}$, where $\mathbf{F}$ is the matrix of derivatives, also known as Jacobi Matrix, $F_{ij} = \delta F_i / \delta r_j$ at $\mathbf{E}(\mathbf{r})$.

If the error of the tetrahedron volume is to be propagated from its 6 edge measurements $r_{12}, r_{13}, r_{14}, r_{23}, r_{24}, r_{34}$, then $n = 6$, $m = 1$ and the variance of the volume can be obtained by

$$\sigma_V^2 = \mathbf{F}\mathbf{C}_\mathbf{r}\mathbf{F}^\mathrm{T}. \tag{3}$$

Assuming uncorrelated range measurements the variance-covariance matrix becomes

$$\mathbf{C}_\mathbf{r} = \operatorname{diag}\left(\sigma_{r12}^2, \sigma_{r13}^2, \sigma_{r14}^2, \sigma_{r23}^2, \sigma_{r24}^2, \sigma_{r34}^2\right), \tag{4}$$

The functional matrix reads

$$\mathbf{F} = \left(\frac{\partial V}{\partial r_{12}} \quad \frac{\partial V}{\partial r_{13}} \quad \frac{\partial V}{\partial r_{14}} \quad \frac{\partial V}{\partial r_{23}} \quad \frac{\partial V}{\partial r_{24}} \quad \frac{\partial V}{\partial r_{34}}\right), \tag{5}$$

and its elements are the Cayley–Menger bideterminants

$$\begin{aligned}
\frac{\partial V}{\partial r_{12}} &= \frac{-1}{36}\frac{r_{12}}{D} D(1, 3, 4; 2, 4, 3), \quad & \frac{\partial V}{\partial r_{13}} &= \frac{-1}{36}\frac{r_{13}}{V} D(1, 2, 4; 2, 3, 4),\\
\frac{\partial V}{\partial r_{14}} &= \frac{-1}{36}\frac{r_{14}}{V} D(1, 2, 3; 3, 2, 4), \quad & \frac{\partial V}{\partial r_{23}} &= \frac{-1}{36}\frac{r_{23}}{V} D(1, 2, 4; 1, 4, 3),\\
\frac{\partial V}{\partial r_{24}} &= \frac{-1}{36}\frac{r_{24}}{V} D(1, 2, 3; 1, 3, 4), \quad & \frac{\partial V}{\partial r_{34}} &= \frac{-1}{36}\frac{r_{34}}{V} D(1, 2, 3; 2, 1, 4).
\end{aligned} \tag{6}$$

The expression $D(s_1; s_2)$ denotes the Cayley–Menger bideterminant, which is a generalisation of the Cayley–Menger determinant (1). The arguments s_1 and s_2 are two dissimilar sequences of points with the length n. Denoting the point

[†]The triangle inequality states that the sum of the lengths of any two sides of a triangle is greater than the length of the remaining side.

numbers $[1, 2, \ldots, n]$ and $[1', 2', \ldots, n']$ the Cayley–Menger bideterminant reads

$$D(1, 2, \ldots, n; 1', 2', \ldots, n') = 2\left(\frac{-1}{2}\right)^n \det\begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & r^2_{11'} & r^2_{12'} & \cdots & r^2_{1n'} \\ 1 & r^2_{21'} & r^2_{22'} & \cdots & r^2_{2n'} \\ \vdots & \vdots & \vdots & \ddots & r^2_{3n'} \\ 1 & r^2_{n1'} & r^2_{n2'} & \cdots & r^2_{nn'} \end{pmatrix}, \tag{7}$$

where r^2_{ij} denote the squared distances between node i and j. The geometric interpretation of the determinant is as follows: if the point sequences s_1 and s_2 (for $n = 3$) form two triangles A_{123} and $\mathrm{A}_{1'2'3'}$ of a tetrahedron with a dihedral angle φ, then $D(1, 2, 3; 1', 2', 3') = 4\,\mathrm{A}_{123}\,\mathrm{A}_{1'2'3'}\cos(\varphi)$ is valid.

Given the volume V, its variance σ^2_V (or its standard deviation σ_V) it is straight forward to determine the probability for a negative value for the volume, which is also the probability of assuming an incorrect topology. Provided that the volume is a normal distributed random variable x with mean V and variance σ^2_V its probability function is

$$P(x) = \frac{1}{\sigma^2_V\sqrt{2\pi}} e^{-(x-V)^2/(2\sigma^2_V)}. \tag{8}$$

The probability that the volume is negative can be determined by the cumulative distribution function, which is the integral of the normal distribution function. It gives the probability that a variate will assume a value smaller than a given limit. With the limit $V_L = 0$ we obtain the probability of a negative volume of

$$P_{V<0} = \int_{-\infty}^{V_L=0} P(x)dx. \tag{9}$$

A one-sided hypothesis test is performed by choosing a confidence limit α, say $\alpha = (1 - 0.997)$. Then, the tetrahedron will be accepted or rejected depending on whether the probability of a negative volume is smaller ($P_{V<0} \leq \alpha$) or higher ($P_{V<0} > \alpha$) than the chosen limit.

If the two tetrahedrons T_{1234} and T_{1235} assumed to be robust according to the volume test, the robustness of the quint can be tested according to the next section.

3.1.4 Ambiguity Test for a Quintilateral

After the two tetrahedrons of a quint have passed the volume test described in the previous section, we can now solve the ambiguity problem of a quint, which was raised in Section 3.1.1.

In order to decide, whether node 5 or its reflected position $5'$ is correct, we perform two two-sided hypothesis tests, namely formulate null hypotheses for the smaller and one for the larger difference between the measured and the

calculated distances $\Delta_{45} = |r_{45} - d_{45}|$ and $\Delta_{45'} = |r_{45} - d_{45'}|$. The quint can only be considered to be robust, if the smaller difference $\Delta_{\min} = \min(\Delta_{45}; \Delta_{45'})$ is a result of pure chance, while the larger difference $\Delta_{\max} = \max(\Delta_{45}; \Delta_{45'})$ clearly differs from zero so that its null hypothesis is rejected.

The two hypothesis tests have three possible outcomes:

1. Both, the null hypothesis for $\Delta_{\min}$ and for $\Delta_{\max}$ are accepted. In this case a clear decision on one of the mirroring positions cannot be made. The quint is considered as not robust and its rejection is recommended.
2. Both null hypotheses are rejected. This case suggests that there is an outlier in one of the 10 range measurements. The quint is to be rejected while the range observations are further tested for outliers.
3. The hypothesis for $\Delta_{\min}$ is accepted and for $\Delta_{\max}$ rejected. This is the only case where the quint is accepted. In case the smaller difference $\Delta_{\min}$ is Δ_{45}, both nodes 4 and 5 lie on the same side of the base plane, if $\Delta_{45'}$ emerges as $\Delta_{\min,}$ the nodes 4 and 5 are on opposite sides. Note that the case where the "hypothesis for $\Delta_{\min}$ is rejected while at the same time that for $\Delta_{\max}$ is accepted" is impossible.

In order to perform these tests, it is necessary to determine the values for Δ_{45}, $\Delta_{45'}$ and their variances $\sigma^2_{\Delta 45}$ and $\sigma^2_{\Delta 45'}$. These 4 quantities are exclusively calculated from the quints' 10 range measurements $r_{12}, r_{13}, r_{14}, r_{15}, r_{23}, r_{24}, r_{25}, r_{34}, r_{35}, r_{45}$ and their variances respectively. The expressions for the 4 quantities are given below.

In order to determine $\Delta_{45} = |r_{45} - d_{45}|$ and $\Delta_{45'} = |r_{45} - d_{45'}|$, it is necessary to derive expressions for the calculated distances d_{45} and $d_{45'}$ without the use of the observation r_{45}. According to Figure 5 the calculated edge lengths between nodes 4 and 5 are:

$$\begin{aligned} d_{45} &= \sqrt{\Delta h^2 + \Delta q^2} = \sqrt{(h_5 - h_4)^2 + \Delta q^2} \\ d_{45'} &= \sqrt{\Delta h'^2 + \Delta q^2} = \sqrt{(h_5 + h_4)^2 + \Delta q^2}. \end{aligned} \tag{10}$$

The heights h_4 and h_5 of the tetrahedrons T_{1234} and T_{1235} can be determined by $h_4 = 3V_{1234}/A_{123}$ and $h_5 = 3V_{1235}/A_{123}$ where V_{1234} and V_{1235} are the tetrahedrons' volumes and A_{123} the area of their common base plane triangle. The volume V of a general tetrahedron is given in (2) and the area of a triangle can be obtained from (1) where dimension $n = 2$. For the area A_{123} we get

$$A^2_{123} = \frac{-1}{16} D(1, 2, 3) = \frac{-1}{16} \det \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & r^2_{12} & r^2_{13} \\ 1 & r^2_{12} & 0 & r^2_{23} \\ 1 & r^2_{13} & r^2_{23} & 0 \end{pmatrix}. \tag{11}$$

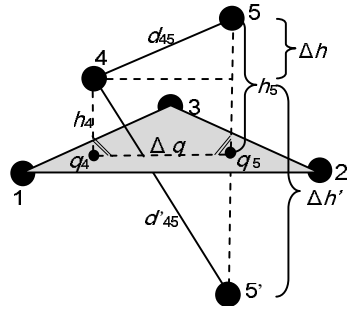

FIGURE 5
Determination of the distance d_{45} and d'_{45} without usage of range measurement r_{45}.

Hence, the squared tetrahedron heights can be written as:

$$h_4^2 = \left(3\frac{V_{1234}}{A_{123}}\right)^2 = -\frac{1}{2}\frac{D(1,2,3,4)}{D(1,2,3)},$$
$$h_5^2 = \left(3\frac{V_{1235}}{A_{123}}\right)^2 = -\frac{1}{2}\frac{D(1,2,3,5)}{D(1,2,3)}. \tag{12}$$

The squared distance Δq^2 between the orthogonal projections of nodes 4 and 5 onto the base plane can be expressed as:

$$\Delta q^2 = 16\big(D_4^2 r_{12} + D_5^2 r_{13} + 2D_4 D_5 D(1,2;1,3)\big)/D^2(1,2,3), \tag{13}$$

where $D_4 = D(1,2,3;1,3,4) - D(1,2,3;1,3,5)$ and $D_5 = D(1,2,3;1,2,4) - D(1,2,3;1,2,5)$ are abbreviations for the differences of Cayley–Menger bideterminants.

Next, we need expressions for the variances $\sigma^2_{\Delta 45}$ and $\sigma^2_{\Delta 45'}$ of the discrepancies Δ_{45} and $\Delta_{45'}$ between the measured and the calculated distances. Given a 10×10 *a priori* covariance matrix of the 10 range measurements that form a quint, the variances for Δ_{45} and $\Delta_{45'}$ can be determined by making use of the law of error propagation.

If a quint has passed the tests on robustness, the assignment of its coordinates can be carried out, which is described in the next section.

3.1.5 Assigning Local Coordinates to a Quint

This step uses inner geometric features such as volumes and edge lengths to assign local Cartesian coordinates to the 5 nodes of a quint. The local reference frame used here is an ordinary minimum constraints network datum. Our choice of a local datum that defines the 3 translations and 3 rotations is as follows: node 1 is set as the point of origin, node 2 on the positive x-axis, node 3 on the x-y-plane and node 4 on a positive z-coordinate axis. The z-coordinate of node 5 is assigned a positive or negative value depending on

the outcome of the ambiguity problem as discussed in the previous section. Note that the allocation of the coordinates for nodes 4 and 5 is identical to a 3D-trilateration from the nodes 1, 2 and 3. The local coordinates read

$$\mathbf{P}_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} r_{12} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{P}_3 = \begin{pmatrix} D(1,2;1,3)/r_{12} \\ \frac{1}{2}\sqrt{-D(1,2,3)}/r_{12} \\ 0 \end{pmatrix},$$

$$\mathbf{P}_4 = \begin{pmatrix} 4(D(1,2,3;1,3,4)r_{12} - D(1,2,3;1,2,4)D(1,2;1,3)/r_{12})/D(1,2,3) \\ 2D(1,2,3;1,2,4)/\left(\sqrt{-D(1,2,3)}r_{12}\right) \\ \sqrt{\frac{-1}{2}D(1,2,3,4)/D(1,2,3)} \end{pmatrix}$$

$$\mathbf{P}_5 = \begin{pmatrix} 4(D(1,2,3;1,3,5)r_{12} - D(1,2,3;1,2,5)D(1,2;1,3)/r_{12})/D(1,2,3) \\ 2D(1,2,3;1,2,5)/(\sqrt{-D(1,2,3)}r_{12}) \\ \pm\sqrt{\frac{-1}{2}D(1,2,3,5)/D(1,2,3)} \end{pmatrix}. \tag{14}$$

This assignment of coordinates does not make use of the redundant observation r_{45}, which means that the determination is consistent and constrains have not been introduced. Consequently, the errors from all range measurements accumulate in the distance d_{45}. Using the unconstrained coordinates as such is not the preferable basis for cluster expansion. Optimally, the measurement errors are distributed evenly over all edges according to the uncertainties. The next paragraph details how the optimal values can be obtained by a least squares adjustment.

3.1.6 Least Squares Adjustment/Refinement

In a Gauss–Markov model the system of linear equations for the parametric least-squares adjustment can be written as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \quad \text{rank}(\mathbf{A}) = n < m, \tag{15}$$

where $\mathbf{y}$ is a $m \times 1$ vector of observations (here: range measurements), $\mathbf{x}$ is the $n \times 1$ vector of unknown parameters (here: coordinates of the nodes), $\mathbf{A}$ is the $m \times n$ design matrix of known coefficients, and e is a normally distributed $m \times 1$ random error vector, where $\mathbf{e} \sim (0, \mathbf{C}_l)$. The stochastic model is given by the covariance matrix $\mathbf{C}_l = \sigma_0^2 P^{-1}$ of the observations, where σ_0^2 is the *a priori* variance factor. $\mathbf{C}_l$ leads to the weighting matrix $\mathbf{P}$. Using the best linear uniformly unbiased estimator, the least squares solution then arises from

$$\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{P}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}\mathbf{y} = \mathbf{N}^{-1}\mathbf{A}^T\mathbf{P}\mathbf{y}, \tag{16}$$

where $\mathbf{N}$ is the matrix of normal equations. The residuals $\mathbf{v}$, the estimated variance component (or reference factor) $\hat{\sigma}_0^2$ and the covariances of the unknowns $\mathbf{C}_x$ become

$$\mathbf{v} = \mathbf{A}\hat{\mathbf{x}} - \mathbf{y}, \quad \hat{\sigma}_0^2 = \mathbf{v}^T\mathbf{P}\mathbf{v}/(m-n) \quad \mathbf{C}_x = \hat{\sigma}_0^2\mathbf{N}^{-1}. \tag{17}$$

Applying the least-squares adjustment to a 3-dimensional distance based network leads to a datum defect of at least $d = 6$, due to 3 translations and 3 rotations, which make up 6 degrees of freedom. As a consequence, the design matrix $\mathbf{A}$ and the normal matrix $\mathbf{N}$ in (16) are singular with a rank defect of at least 6, prohibiting an inversion. This problem can be solved by applying a network datum i.e. assignment of approximate coordinates to all points in the network as it is done exemplarily in Section 3.1.5. However, by defining fixed values to some coordinates (e.g. $\mathbf{P}_1 = (0, 0, 0)$), their corresponding variances are zero. Such a datum definition, also called "zero-variance computational base", does not put constraints to the inner geometry of the network. Consequently, this approach is known as free network adjustment. It is obvious that any choice for a particular datum definition leads to a solution vector $\hat{\mathbf{x}}$ and a covariance matrix of the unknowns $\mathbf{C}_x$ which cannot be used to assess a network's quality. For example, the variances of $\mathbf{P}_1$ in (14) will be zero, while those for $\mathbf{P}_4$ and $\mathbf{P}_5$ may be overestimated. Unrealistic variances may cause problems when other points are to be added to the network.

Although coordinates are in general non-estimable parameters of a network, we can minimise the impact of datum assignment by applying datum constraints in such a way, that all points in the network participate equally in the definition of the zero-variance elements by demanding $\text{trace}(\mathbf{Q}_{xx}) = \min$. In a 3D distance based network one realisation is to demand that:

a) the centroid of the adjusted coordinates has a zero variance and takes the same value as the centroid of the initial approximate coordinates and,

b) the overall rotation of all adjusted coordinates versus their initial approximate values is zero with a zero variance.

This so called inner constraints datum can be applied by extending the system of normal equations with a matrix $\mathbf{G}$, where

$$\mathbf{G}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & \cdots & 0 & 0 & 1 \\ 0 & z_1^0 & -y_1^0 & 0 & z_2^0 & -y_2^0 & \cdots & 0 & z_p^0 & -y_p^0 \\ -z_1^0 & 0 & x_1^0 & -z_2^0 & 0 & x_2^0 & \cdots & -z_p^0 & 0 & x_p^0 \\ y_1^0 & -x_1^0 & 0 & y_2^0 & -x_2^0 & 0 & \cdots & y_p^0 & -x_p^0 & 0 \end{pmatrix}. \tag{18}$$

The extended system of normal equations

$$\begin{pmatrix} \mathbf{N} & \mathbf{G} \\ \mathbf{G}^{\mathrm{T}} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{x}} \\ \mathbf{k} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{\mathrm{T}}\mathbf{P}\mathbf{y} \\ \mathbf{0} \end{pmatrix}, \tag{19}$$

with correlate vector **k**, is invertible and its solution presents the most homogeneous spread of random error through the network.

For a free least squares adjustment of a quint there are $n = 15$ coordinate unknowns, $m = 10$ range observations and a datum defect of $d = 6$. With a redundancy $r = m - n + d = 1$ the estimated variances depend only on a single inconsistency and are consequently not to be considered as reliable. However, as the quint is extended by lateration, the redundancy in the network increases throughout. In a 'richly redundant' network the estimated variances become a meaningful measure of the coordinate uncertainties.

3.2 Expansion of the Smallest Rigid Structure

While Section 3.1 elaborated the construction of a quint as the smallest rigid cluster, this section suggests an expansion scheme that incrementally adds neighbouring nodes to a cluster. Since the multilateration technique we use differs from the traditional approach, we first explain multilateration in general (Section 3.2.1), and then show how multilateration is conventionally carried out in Section 3.2.2 followed by our robust approach in Section 3.2.3.

3.2.1 Least Squares Multilateration

A key requirement of our approach is that adding nodes to a cluster must be carried with redundancy in mind. Considering the method of lateration, any prospect candidate for becoming a member of the cluster requires to share at least 4 ranges with the cluster. If 4 ranges are present (in 3 dimensions), the minimum redundancy is $r = 1$ and satisfies our requirement of over determination. Figure 6a shows an example of multilateration of a node A from 4 anchor nodes to a larger cluster. These 4 anchor nodes themselves do not necessarily need to share ranges to each other as illustrated in Figure 6b. Consequently, demanding that the nodes 1, 2, 3, 4 and A make up a fully connected quint is an unreasonable constraint, that restricts the localisation of nodes unnecessarily. Moore *et al*. (2004) however follow such a strict approach in 2D, when demanding that only nodes of robust quadrilaterals can be attached to a cluster. In 3D, such a restrictive approach would correspond to solely allocating

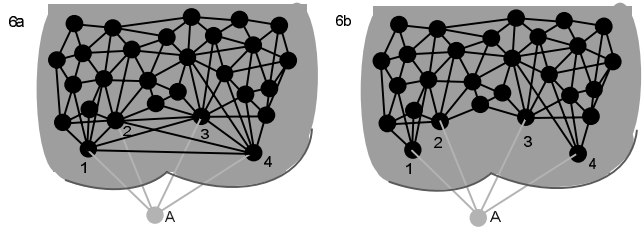

FIGURE 6
Multilateration of node A to a cluster from 4 cluster nodes. In 6a the added point belongs to a quint; in 6b this is not the case.

nodes that are part of a quint. For this reason we do not follow an approach that finds all possible quints and then merges them into one big cluster using transformations.

Our robust multilateration method exclusively uses range observations and is not based on distance bounding as in Čapkun and Hubaux (2004). Distance bounding is a technique by which a node can obtain an upper-bound on its distance to another node.

3.2.2 Traditional Multilateration

Our approach in lateration breaks away from the established method for multilateration as proposed in Gober *et al.* (2005), Langendoen and Reijers (2003), Savvides *et al.* (2001) Niculescu and Nath (2001) and Beutel (1999). Due to its widespread use, we repeat the traditional method here.

The common multilateration method for determining the position of an object based on simultaneous range measurements from more than 3 stations located at known sites starts off with the system of observation equations

$$r_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}, \tag{20}$$

where $\mathbf{P}_i = (x_i, y_i, z_i)$, $i = 1, 2, \ldots, n$ are the known coordinates of station i, and r_i is the range measurement associated with it. The unknown position is denoted by $\mathbf{P}(x, y, z)$. This problem is equivalent to finding the intersection point(s) of n spheres in $\mathbb{R}^3$ and is conventionally solved by squaring the equations (20) and subtracting the last equation from the first $n-1$ equations. These algebraic manipulations lead to the system of linear equations

$$\begin{aligned}
&x_1^2 - x_n^2 - 2(x_1 - x_n)x + y_1^2 - y_n^2 - 2(y_1 - y_n)y \\
&\quad + z_1^2 - z_n^2 - 2(z_1 - z_n)z = r_1^2 - r_n^2 \\
&\qquad\qquad\qquad\qquad \vdots \\
&x_{n-1}^2 - x_n^2 - 2(x_{n-1} - x_n)x + y_{n-1}^2 - y_n^2 - 2(y_{n-1} - y_n)y \\
&\quad + z_{n-1}^2 - z_n^2 - 2(z_{n-1} - z_n)z = r_{n-1}^2 - r_n^2.
\end{aligned} \tag{21}$$

The terms are recorded in the form $\mathbf{Ax} = \mathbf{y}$, where

$$\begin{aligned}
\mathbf{A} &= \begin{pmatrix} 2(x_1 - x_n) & 2(y_1 - y_n) & 2(z_1 - z_n) \\ \vdots & \vdots & \vdots \\ 2(x_{n-1} - x_n) & 2(y_{n-1} - y_n) & 2(z_{n-1} - z_n) \end{pmatrix}, \\
\mathbf{y} &= \begin{pmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 + z_1^2 - z_n^2 + r_n^2 - r_1^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 + z_{n-1}^2 - z_n^2 + r_n^2 - r_{n-1}^2 \end{pmatrix}.
\end{aligned} \tag{22}$$

The linear system is solved for $\mathbf{P}(x, y, z) = \hat{\mathbf{x}}$ using $\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{y}$ as given in (16). The pivotal point about this conventional lateration approach is the fact that $\hat{\mathbf{x}}$ is not a least-squares estimator for the non-linear system (20). In a proper LS approach the L_2-Norm is minimised

$$\sum_{i=1}^{\mathrm{n}} e_i^2 = \min, \tag{23}$$

and the observation equations read

$$E\{r_i\} = r_i - e_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}, \tag{24}$$

where E is denoting the expectation operator and the random errors e_i are assumed to be distributed normally. Squaring the observation equations would lead to non-linear terms for the random errors

$$r_i^2 + e_i^2 - 2r_i e_i = (x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2. \tag{25}$$

In the conventional lateration approach however, these non-linear terms are neglected, leading to a different problem other than the LS minimisation (23). Consequently, solving the simplified problem does not yield the LS solution. The conventional lateration method can fail even for a system with tiny inconsistencies. In addition to the problem of linearisation, the conventional multilateration method is incompatible with correctly solving the mirroring problem. In our opinion, this is the reason, why iterative multilateration is considered as unreliable in the presence of random errors. In a large network a single incorrect placement of a node may cause a failure of the whole localisation algorithm where lateration is carried out successively. We argue that multilateration can be carried out reliably, if the approach described in the next section is followed.

3.2.3 A Robust Approach for Multilateration

The main phases of the process to attach an additional point to a cluster by multilateration are: 1) trilateration 2) test on geometric robustness 3) solving ambiguity 4) LS adjustment of the new point 5) network adjustment 6) post-adjustment analyses. Each phase is detailed below.

1) Out of $n > 3$ known stations 3 are chosen arbitrarily, say $\mathbf{P_1}$, $\mathbf{P_2}$ and $\mathbf{P_3}$, from where the new point $\mathbf{P}(x, y, z)$ is determined by 3D trilateration. A straight forward formulation for trilateration introduced by Thomas and Ros (2005) reads

$$\mathbf{P}, \mathbf{P}' = \mathbf{P}_1 + 4\Big(D(1,2,3;1,3,\mathrm{p})[\mathbf{P_2} - \mathbf{P}_1] - D(1,2,3;1,2,\mathrm{p})[\mathbf{P}_3 - \mathbf{P}_1] \mp \sqrt{\tfrac{1}{8}D(1,2,3,\mathrm{p})}[\mathbf{P_2} - \mathbf{P}_1] \times [\mathbf{P}_3 - \mathbf{P}_1]\Big) \Big/ D(1,2,3), \tag{26}$$

where $D(1, 2, 3)$ and $D(1, 2, 3, \mathrm{p})$ are the Cayley–Menger determinants given in (1). $D(1, 2, 3; 1, 3, \mathrm{p})$ and $D(1, 2, 3; 1, 2, \mathrm{p})$ are the Cayley–Menger

bideterminants according to (7). Here, the entries for the Cayley–Menger determinants are the distances between the known stations r_{12}, r_{13} and r_{23} that are determined from coordinates and r_{1p}, r_{2p}, r_{3p} are the measured distances to the new point.

2) The robustness of the tetrahedron T_{123p} spanned by $\mathbf{P_1}$, $\mathbf{P_2}$, $\mathbf{P_3}$ and $\mathbf{P}$ is tested following the procedure described in Section 3.1.3. Variances of ranges that have not been measured can be estimated from the positional errors by covariance propagation. The rejection or acceptance of this point-combination depends on the geometry and the error variances. In case a configuration is rejected, the algorithm continues with another selection of known stations.

3) A decision has to be made whether point $\mathbf{P}$ or its reflection $\mathbf{P}'$ (which is mirrored at the plane spanned by $\mathbf{P_1}$, $\mathbf{P_2}$ and $\mathbf{P_3}$) is correct. Therefore, we chose another known station $\mathbf{P_4}$ and compute all 10 range measurements and their covariances while incorporating the uncertainty of the position estimates by error propagation. This approach follows the ambiguity test described in Section 3.1.4, where the notation can be adopted here by denoting $\mathbf{P}$ as $\mathbf{P_5}$. The ambiguity test will either result in a rejection of both points ($\mathbf{P}$ and $\mathbf{P}'$), if the ambiguity problem could not be solved definitely or the acceptance of either. In case of a rejection the test is repeated using another set of known stations. If all combinations fail, which is the case if the stations are close to one plane or form other close-to-singular formations, the lateration of $\mathbf{P}$ to the cluster has failed and is abandoned.

4) The coordinates for $\mathbf{P}$ are refined following the least squares adjustment described in Section 3.1.6. Here, the initial approximate coordinate values for $\mathbf{P}$ are denoted by $\mathbf{P_0}(x_0, y_0, z_0)$. The observation equations (24) are linearised and the terms are recorded in

$$\mathbf{A} = \begin{pmatrix} \frac{x_0 - x_1}{d_1} & \frac{y_0 - y_1}{d_1} & \frac{z_0 - z_1}{d_1} \\ \vdots & \vdots & \vdots \\ \frac{x_0 - x_n}{d_n} & \frac{y_0 - y_n}{d_n} & \frac{z_0 - z_n}{d_n} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} r_1 - d_1 \\ \vdots \\ r_n - d_n \end{pmatrix}, \tag{27}$$

where r_i are the observed distances and d_i the calculated distances between known station i and the approximate position $\mathbf{P_0}(x_0, y_0, z_0)$. Assuming uncorrelated distance measurements, the weight matrix is $\mathbf{P} = \mathrm{diag}(\sigma_0^2/\sigma_1^2, \ldots, \sigma_0^2/\sigma_n^2)$, where σ_0^2 is the *a priori* variance factor and σ_i^2 the *a priori* variance of r_i. Now, the adjustment can be carried out using (16) and (17). It is important to iterate (16) and (17) with updated approximate values for $\mathbf{P_0}$ until

a) further improvements for $\hat{\mathbf{x}} = \mathbf{P_0}(x_0, y_0, z_0)$ are negligible and

b) the estimated variance component $\hat{\sigma}_0^2$ is not decreasing or

c) the upper limit of iterations has been reached (in order to bound computational costs).

5) The whole cluster is adjusted. Since a network adjustment of a larger cluster can be prohibitive in terms of computational costs the decision whether to adjust the whole network or not is a trade off between benefit and computation time. While this step can be omitted after one single node has been added to the cluster, the adjustment is an obligatory step at the end of the clustering phase in order to obtain the complete covariance matrix $\mathbf{C}_x$. The advantage of an adjustment is that the constraints and the errors that have been introduced by the new point **P** are distributed evenly (in a LS-sense) in the network. As a consequence, the accumulation of errors is avoided when further nodes are laterated to the cluster – in particular because additional points tend to be laterated from new cluster members lying on the outskirts of the cluster. The free network adjustment is carried out as described in Section 3.1.6 using the inner constraints datum (18).

6) Post-adjustment analyses are carried out to provide a statement on the quality of the estimates. Various statistical techniques such as global test of the model and tests for blunders and outliers are given in Anderson and Mikhail (1988). If the statistical tests on the residuals fail, the solution is rejected and the measurements are searched for outlier observations.

After the statistical tests have been passed successfully, the coordinate estimates for the point **P** are assumed to be 'robust'. **P** becomes a member of the cluster with known station coordinates $\mathbf{P}(x, y, z)$ and an associated 3×3 covariance matrix $\mathbf{C}_x$, which can be obtained from the large $3n \times 3n$ covariance matrix as a result of step 5. Note that only storing the 3×3 covariance matrices is a simplification of the error model that neglects the inter-node correlations.

The next section outlines how the repetition of the multilateration technique described in this section is a means for the creation of large clusters.

3.2.4 Iterative Collaborative Multilateration

The proposed clusterisation strategy can be classified as *multilateration* because each point is attached to the cluster using multiple (4 or more) range measurements to known cluster-points according to Section 3.2.3. The method is of iterative type because nodes are added to the cluster sequentially and becoming known anchor points for further expansion of the cluster. Finally, our algorithm is *collaborative* because nodes may also be added jointly if the requirements for multilateration are not satisfied. The example illustrated in Figure 7 shows that nodes with less than 4 range observations to a cluster can be uniquely located, if additional ranges to non-localised nodes are present. In case of the given example, nodes A and B are connected to the cluster by 3 ranges each. Furthermore, the distance between A and B has been observed.

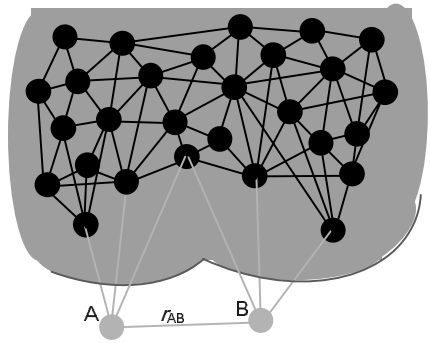

FIGURE 7
Collaborative multilateration of nodes A and B.

If A had a unique position, B could be uniquely located as well, and vice versa. The solution is obtained by sticking to the multilateration approach in Section 3.2.3 with the only dissimilarity that r_{AB} disambiguates between 4 points A, A′, B and B′. The localisation of A and B is still over-determined and therefore, in accordance with our philosophy on redundancy at every step.

More complicated cases exist, where several nodes can be attached to a cluster collaboratively. The theory of collaborative subtrees for detection of unique possible solutions has been well developed by Aspnes *et al.* (2004). However, one has to consider computational costs and the fact that theoretically uniquely localisable subtrees turn out to be non-localisable in practice due to the presence of measurement noise.

The next section describes how clusters obtained by iterative multilateration can be merged to one single cluster.

3.3 Merging of Clusters

The growth of a cluster using multilateration is concluded when no more nodes can be found that are eligible for lateration. On the other hand, the remaining nodes are likely to make up their own clusters. If we think of a network in a building, a cluster may be set up in one room while the lower density of devices in corridors prohibits an extension of the cluster to neighbouring rooms. If different clusters exist in various rooms, the question is, whether they can be combined. For the localisation algorithm, there is the fundamental need for a merging capability of clusters that are somehow linked.

Two clusters are merged by shifting and rotating one coordinate system into the other. If two clusters share at least 4 nodes they can be merged by the means of an over-determined 6-parameter transformation. Finding the relationship between two coordinate systems is a least-squares problem for three or more control points. We propose the closed-form solution developed by Horn (1987), which is robust and bounded in terms of computational costs.

The 3 shift parameters are computed by determining the centroids of both coordinate sets. The problem of finding the 3 rotation parameters is formulated as an eigenvalue problem of a symmetric 4×4 matrix. To find the eigenvalues, a quartic equation with one coefficient being zero has to be solved. For non-aligned point formations, this problem has a unique solution. Since the transformation is carried out with redundancy, the assessment of this step can be controlled by the residuals. Subsequent to the 6-parameter transformation the merged cluster can be adjusted for the purpose of a homogenous distribution of the inconsistencies.

Alternatively, if the two clusters share less than 4 nodes but are linked by additional range measurements the 6 parameters for the 3D-transformation are also computable. For an over-determined computation of the transformation parameters with 3 common points, at least one range must be present. The cases with two, one or no control point require at least 2, 4 or 7 ranges respectively. The next section describes how nodes that don't fulfil the requirements for multilateration can be added to the cluster.

3.4 Coarse Positioning

Not all nodes will be able to participate in the procedure described above. Either, because the number of range observations to neighbouring nodes is not sufficient, or due to badly constrained configurations, single nodes cannot participate in the precise lateration network. Possible situations include scenarios where the complete network adjustment fails. However, users do demand positioning for all nodes even under these circumstances.

To enable users to have a level of positioning capability in very difficult environments, a coarse positioning service is introduced for cases where the precise mode fails. Coarse positioning exploits connectivity information between nodes when sufficient range measurements are not available. If a node is only connected to a single fixed neighbour, its position is enclosed within a sphere of a radius being as large as the range itself. In a straightforward way, the new node's coordinates are assigned to be identical to its fixed neighbour, but its position errors are set to the actual range value.

In a similar fashion, the cases with 2 or 3 available range measurements are dealt with. For example, if 3 ranges to known nodes are available, the new point can be trilaterated from those nodes. However, since it is not clear which of the two embeddings is correct and there is no 4th range available to disambiguate between them, the best we can do is to locate the new node between the two candidate points with an attached error bar of half the distance between these two embeddings. Clearly, the reliability measures for these nodes need to be set as 'very low' and participation in the actual geodetic network is not possible. Future research will also explore ways of reducing the instances when coarse positioning is invoked through the use of historical data, map-matching, etc.

The next section describes how the merged cluster with local coordinates can be transformed into a reference system, if anchor nodes are present.

3.5 Transformation into a Reference System

The last step in the localisation algorithm is to transform the local coordinates of the cluster into a reference system. The requirement for a successful transformation is that at least 3 points of the cluster are control points in the target system. Analogous to the merging of clusters in Section 3.3, a 6-parameter transformation is carried out. Note that the transformation does not solve the mirroring ambiguity problem. Since the reference system is definite in its embedding but the local coordinates of a generated cluster have an arbitrary reflection the correct reflection must be selected. The correct reflection is discovered simply by comparing the residuals of the transformation.

Finally, a fully constrained LS network adjustment is performed that combines all available anchor nodes and all range measurements and refines all approximate positions simultaneously. The mean errors of the coordinates and the point confidence ellipses are taken from the variance-covariance matrix.

4 SIMULATION RESULTS

In order to confirm the proposed localisation theory and to predict the performance of the positioning system currently under development (iPLOT), the results of a number of simulations are given in this section.

4.1 Evaluation Criteria

First, we look at the size of the biggest cluster containing anchor nodes. An important criterion of the measurement noise level is the mean deviation (rms) of the ranges

$$\sigma_r = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(r_i - d_i)^2}, \tag{28}$$

where m is the number of the range observations r_i, and d_i are the true distances. In the simulations the observations r_i are generated by a normal distribution function with mean d_i and standard deviation σ_r according to the expectation function $E\{r_i\} = N(d_i, \sigma_r)$. It is useful to compare σ_r with the mean deviation σ_p (rms) of the positions

$$\sigma_p = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\mathbf{P}}_i - \mathbf{P}_i)^2}, \tag{29}$$

where n is the number of nodes, $\mathbf{P}_i$ the true position vector and $\hat{\mathbf{P}}_i$ the estimated position vector of the localised node i. A criterion that is less sensitive to outliers is the average position deviation

$$a_p = \frac{1}{n}\sum_{i=1}^{n}\sqrt{(\hat{\mathbf{P}}_i - \mathbf{P}_i)^2}. \tag{30}$$

In order to compare the true and the estimated deviations, the average of the estimated mean square positional errors

$$\hat{a}_p = \frac{1}{n} \sum_{i=1}^{n} \sqrt{q_{xxi} + q_{yyi} + q_{zzi}}, \tag{31}$$

is determined, where q_{xxi}, q_{yyi}, q_{zzi} are diagonal elements of the variance-covariance matrix $\mathbf{C}_x$ as a result of the network adjustment. The internal consistency of the free network can be assessed by the square root of the estimated reference variance

$$\hat{\sigma}_r = \sqrt{\frac{1}{m-3n} \mathbf{v}^{\mathrm{T}} \mathbf{v}} = \sqrt{\frac{1}{m-3n} \sum_{i=1}^{m} (\hat{r}_i - r_i)^2}, \tag{32}$$

where $\mathbf{v}$ is the vector of residuals containing the differences between the estimated distances $\hat{r_i}$ (obtained from a LS-adjustment) and the measured distances r_i. If the model is correct, estimator $\hat{\sigma}_r$ should be found equal to σ_r using a hypothesis test of the X^2 distribution with $m - 3n$ degrees of freedom.

4.2 Simulation Setup

In order to asses the impact of noise, we set up a simple scenario that does not address other real-world problems such as the none-line-of-sight or multipath propagation. We place 100 nodes uniformly and randomly in a three-dimensional cube with 100 m edge length. The range length is recorded only between nodes that are within a maximum ranging distance $r_{\max}$. For a later comparison, the positions are recorded as true positions $\mathbf{P}_i$ and error free ranges d_i. Furthermore, 10 nodes are randomly selected to serve as control points and saved separately. Then normally distributed random noise with mean $\mu = 0$ and standard deviation σ_r is added to the observations, now recorded as $r_i = d_i + \sigma_r z_i$, where z is a function that generates normally distributed random numbers with a standard deviation of 1. Figure 8 shows a sample network.

Based on the methodology described in Section 3, the estimated positions $\hat{\mathbf{P}}_i$ are retrieved. The algorithm only uses the erroneous ranges r_i and, in the last step, the 10 control points to estimate the true positions. The parameters for performance assessment are determined according to Section 4.1.

4.3 Simulation Results

When evaluating the algorithm's performance we are interested in how both node connectivity and measurement noise affect the results. The node connectivity was controlled by the maximal signal range $r_{\max}$.

Figure 9 (with 450 networks of 100 nodes each) shows how the number of localised nodes increases with the connectivity. In case of zero noise the fine localisation mostly takes place with 5 to 10 ranges per node. At $\sigma_r = 1$ m

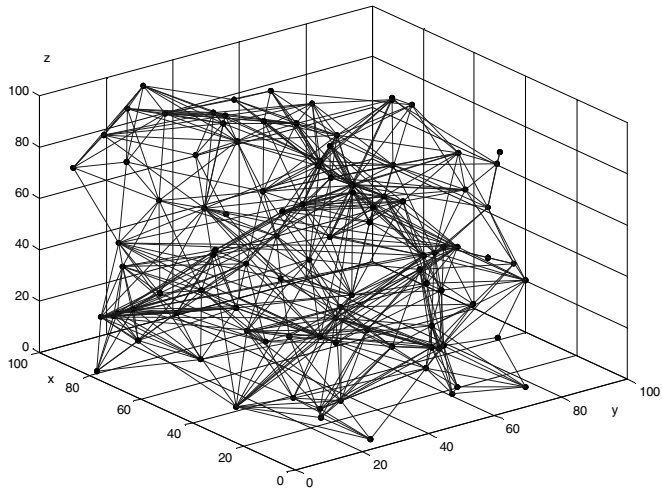

FIGURE 8
Distance network with 100 randomly deployed nodes in a 100 m^3 test cube. The number of ranges is 598 with $r_{max} = 35$ m and an average range length of 28.3 m.

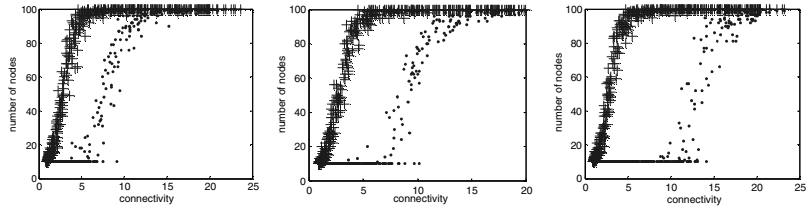

FIGURE 9
Number of localised nodes versus connectivity for the cases with 0 m, 1 m and 5 m mean deviation σ_r in the range measurements. Each dot represents fine localisation and each plus coarse localisation respectively.

(3.5% of the average range length) the localisation success rises to between 8 and 12, at $\sigma_r = 5$ m around 10 to 15 connections per node. The success rate of coarse localisation is independent from the measurement noise, requiring 2 to 4 connections on average.

Figure 10 shows the performance of 450 networks at $\sigma_r = 1$ m noise level for different levels of connectivity. At a connectivity of 15, the mean position deviations σ_p reach an average value of 1 m. At higher connectivity, the deviations from truth are even smaller than the ranging accuracy, reaching $\sigma_p = 0.75$ m. The estimated deviations $\hat{a}_p$ (as a result of the adjustment) are shown with dots. As they are not far from the true deviations, they provide realistic figures for the uncertainties of the localised nodes.

Figure 11 shows how the noise level in the ranges σ_r influences the position deviation a_p of the localised nodes at a maximum signal range $r_{max} = 35$ m.

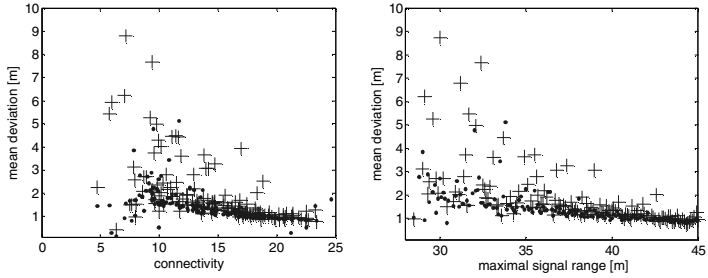

FIGURE 10
True mean position deviations σ_p (pluses) and estimated deviations $\hat{a}_p$ (dots) in [m] for different levels of connectivity at $\sigma_r = 1$ m error budget in the range measurements. 6 outliers > 10 m are not visible in the plots.

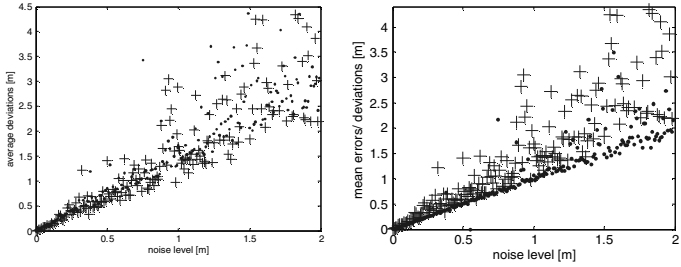

FIGURE 11
True position deviations a_p (pluses) in [m] for measurement noise σ_r between 0 m and 2 m. For comparison, the estimated deviations $\hat{a}_p$ and the estimated reference variances $\hat{\sigma}_r$ are shown (dots) in Fig. 11a and Fig. 11b respectively.

The linear dependencies on σ_r and a_p in Figure 11 have an average proportionality factor of 1. This testifies that our localisation algorithm has almost reached the best possible performance level, which has been determined by Savvides (2003a) as the Cramer Rao Bound behaviour. This is especially true for noise levels with less than 3.5% error (1 m level), where almost all nodes in all networks have been localised correctly. Note that a single falsely located node (e.g. due to a false folding ambiguity) causes the average position deviation to rise significantly. In another experiment, we left out the error propagation for ambiguity assessment and replaced it by a simple strategy that always accepts the solution $\Delta_{\min}$ (see Section 3.1.4 for details). In a similar way, the ambiguity decision is made in the traditional multilateration algorithm (see Section 3.2.2). Figure 12 shows that only a few networks reach the optimal performance level. Figures for different noise levels are given in Table 1.

4.4 Performance on Real Data

Outlier detection is an essential component of the system proposed. Large outliers – such as those resulting from multipath – have not been considered.

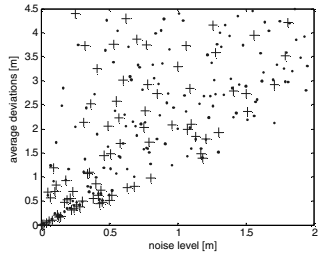

FIGURE 12
True position deviations a_p (pluses) and estimated deviations $\hat{a}_p$ (dots) in [m] without ambiguity control.

σ_r in [m]	0.0	0.1	0.5	1.0	2.0
σ_r in [%] of range	0.0	0.3	1.7	3.5	7.1
localised nodes	99.3	98.8	98.4	98.2 (99.2)	96.0
σ_p in [m]	0.0	0.13	0.75	1.6 (9.2)	4.7
a_p in [m]	0.0	0.10	0.52	1.1 (4.8)	3.0
$\hat{a}_p$ in [m]	0.0	0.11	0.55	1.1 (1.9)	2.5
$\hat{\sigma}_r$ in [m]	0.0	0.10	0.50	1.0 (1.7)	2.1

TABLE 1
Comparison of the localisation performance of 450 networks with 100 nodes each and $r_{\max} = 40$ m at different noise levels. Results without hypothesis tests are given in brackets.

However, the simulation included relatively large white-Gaussian noise which is a good starting point for exposing the performance of the algorithm. Outliers can be detected more easily than smaller errors by following the proposed concept that inconsistent structures do not get connected. Consequently the algorithm is projected to perform well in real environments. Most outlier observations can be eliminated in a prior data filtering (e.g. using symmetric distance enforcement and triangular inequality enforcement, Zhang, 2004). Future research will include the combined presence of noise and outliers.

5 CONCLUSIONS AND FUTURE WORK

The results given in Section 4 show that the current algorithm is performing well without the necessity of backtracking. Our simulations have demonstrated the importance of carefully making decisions on folding ambiguities while creating rigid structures and expanding them by multilateration. We have studied networks with relatively large errors of up to 7.5% of the true ranges and shown that it is possible to achieve a position deviation that is of the size of the ranging error. In case of high inter-node connectivity above 15,

the position deviations can be even smaller than the noise level of the range measurements. In real environments however, this level of performance is not likely to be achieved.

We have also shown that multilateration can be a reliable method of clusterisation. The accumulation of errors can be avoided when the traditional simplified formulation for multilateration is replaced by determining the least squares solution in combination with thorough handling of the fold-ambiguities that arise from uncertainties in the ranges.

Future work will focus on the implementation of the proposed localisation algorithm on real sensors. As it has been suggested from real-world experiments, the error behaviour at the sensors differs from being Gaussian. Moving away from laboratory conditions further challenges will appear, such as constraints on computation time, message routing, multipath (causing outlier observations) and movement of nodes. Future research will also explore ways of reducing the instances when coarse positioning is invoked through the use of historical data, map-matching and mobile-assisted localisation.

REFERENCES

Anderson, J and Mikhail, E (1988): Surveying, Theory and Practice, 7th Edition 1998 McGraw-Hill.

Aspnes, J; Eren, T; Goldenberg, D; Morse, A; Whiteley, W; Yang, Y; Anderson, B and Belhumeur B (2004): A theory of network localization. Submitted to IEEE Transactions on Mobile Computing, July 2004.

Beutel, J (1999) Geolocation in a picoradio environment, Master's thesis, UC Berkeley, Department of Electrical Engineering and Computer Science, 120p.

Čapkun, S and Hubaux, JP (2004): Securing Position and Distance Verification in Wireless Networks. Technical Report IC/200443, EPFL, May 2004.

Gober, P; Ziviani, A; Todorova, P; Dias de Amorim, M; Hunerberg, P and Fdida, S (2005): Topology Control and Localization in Wireless Ad Hoc and Sensor Networks, Ad Hoc & Sensor Wireless Networks, Vol. 1, pp. 301–321

Goldenberg D; Krishnamurthy, A; Maness, W; Yang, Y; Young, A; Morse, A; Savvides, A and Anderson B (2005): Network Localization in Partially Localizable Networks. In Proceedings of IEEE INFOCOM 2005, Miami, FL, March 13–17, 2005.

Grafarend, E and Schaffrin, B (1976): Equivalence of estimable quantities and invariants in geodetic net-works. Zeitschrift für Vermessungswesen 101: pp. 485–491.

Horn, B (1987): Closed-from solution of absolute. orientation using unit quaternions. Journal of Opt. Soc. Amer., vol. A-4, pp. 629–642, 1987.

Langendoen, K and Reijers, N (2003): Distributed localization in wireless sensor networks: a quantitative comparison. The International Journal of Computer and Telecommunication Networking, 43(4):499–518, November 2003.

Manolakis, D (1996): Efficient Solution and Performance Analysis of 3-D Position Estimation by Trilateration, IEEE Aerospace and electronic systems, Vol 32, No. 4, Oct. 1996.

Mautz, R; Ochieng, W Y; Walsh, D; Brodin, G; Cooper, J; Kemp and Lee, T S (2006), "Low-cost intelligent pervasive location and tracking (iPLOT) for the management of crime", Journal of Navigation, vol. 59, no. 2, (2006).

Moore, D; Leonard, J; Rus, D and Teller, S (2004): Robust distributed network localization with noisy range measurements, Proceedings of the ACM Symposium on Networked Embedded Systems, 2004.

Neitzel, F (2003) Identifizierung konsistenter Datengruppen am Beispiel der Kongruenzuntersuchung geodätischer Netze. German Geodetic Commission, Vol. C, No. 565, München.

Niculescu, D and Nath, B (2001): Ad-hoc positioning system, in: IEEE GlobeCom, 2001.

Savarese, C; Langendoen, K and Rabaey, J (2002): Robust positioning algorithms for distributed ad-hoc wireless sensor networks, in: USENIX Technical Annual Conference, Monterey, CA, 2002, pp. 317–328.

Savvides, A; Han, C and Strivastava, M (2001): Dynamic Fine-Grained Localization in Ad-hoc Networks of Sensors, Proceedings of ACM SIGMOBILE 2001, Rome, Italy, July 2001.

Savvides, A; Park, H and Srivastava, M (2003): The n-Hop Multilateration Primitive for Node Localization Problems. MONET 8(4): 443–451.

Savvides, A; Garber, W; Adlakha, S; Moses, R and Srivastava, M (2003a): "On the Error Characteristics of Multihop Node Localization in Ad-Hoc Sensor Networks," Proceedings of the Second International Workshop on Information Processing in Sensor Networks (IPSN'03), Palo Alto, California, 317–332.

Shang, Y; Ruml, W; Zhang, Y and Fromherz, M (2004): Localization from Connectivity in Sensor Networks, IEEE Transactions on Parallel and Distributed Systems, vol. 15, no. 11, pp. 961–974, Nov. 2004.

Thomas, F and Ros, L (2005): Revisiting Trilateration for Robot Localization, IEEE Transactions on Robotics, Vol. 21, No. 1, pp. 93–101, February 2005.

Whitehouse, K; Karlof, C; Woo, A; Jiang, F and Culler, D (2005): The Effects of Ranging Noise on Multihop Localization: An Empirical Study, in Proc. of IPSN, Los Angeles, CA, April, 2005.

Zhang, Y; Yim, M; Ackerson, L; Duff, D and Eldershaw, A (2004): "STAM: A System of Tracking and Mapping in Real Environments", IEEE Wireless Magazine, Dec. 2004.