Quality of Location: Estimation, System Integration and Application

Martin Berchtold, Till Riedel, Christian Decker, Christian Bittel
TecO, University of Karlsruhe
Vincenz-Priessnitz-Str. 3, 76131 Karlsruhe, Germany

Michael Beigl, Monty Beuster
Distributed and Ubiquitous Systems
Muehlenpfordtstr. 23, 38106 Braunschweig, Germany

Abstract*—Accurate location measurement is an important research topic in Pervasive Computing systems and applications. To achieve high performance measurements, the knowledge of the quality of a measurement, a sensor cue, or an inferred location value is required. This paper presents a novel approach in the deliverance of an independent, unified Quality of Location (QoL) value for Location systems. The proposed approach is highly flexible, independent of technology and location inference mechanisms and approaches, integrable into any existing location system, and does neither require knowledge of sensor, nor of application characteristics. The paper proposes both a method to retrieve QoL for a given system, and shows its application in a setting using a simple ultrasound location system. Retrieving QoL requires a multi step process including a unsupervised subtractive clustering method for initial learning, and a supervised network based fuzzy inference systems (ANFIS) for refinement of the parameters. The approach described can be used in settings using heterogeneous systems, devices, and sensors. It is also usable at any abstraction layer and is able to run on small sensor node devices. Technical foundations of the algorithms are an adaptive network based fuzzy inference systems (ANFIS). In this paper we will show the technical principles, its application and evaluate the performance of the system.

I. Introduction

Location has been an important contextual information for Ubiquitous Computing Systems. Location information measurement is based on a large variety of approaches and methods, such as ultrasound time-of-flight, RF signal strength, massive RFID deployment, UWB measurements, or GPS readings. In many Ubiquitous Computing settings an opportunistic approach is used, e.g. a mix of different simple, inexpensive, and heterogeneous technologies. This leads to the often unknown and mostly low quality of the resulting measurements. A wide range of approaches have been developed to take counter measures against the negative impact, when inferring location from error-prone measurements. One way to address this problem is the multi-sensor fusion, other methods use filtering techniques, and many a combination thereof.

This paper addresses an important problem which is placed before processing of location information through fusion or filtering: estimating the quality of the measurements. Our approach will return a unified, thus modality, algorithm, and sensor fusion stage independent value, describing the quality of any measurement value. Such a quality measure has several applications in Ubiquitous Computing settings, and we will show how it can be used to improve location estimation in real-world settings.

A simple, but typical example setting should make the potential more clear. Figure 1 shows several objects in an environment, which are able to measure the distance to other devices in vicinity. It is desired to select the best measurement alternative, using a comparable quality metric. For instance, when reading the distance from A to B, either the direct measurement or the indirect measurement A-C-B of distances could be used. Due to lower quality in short-distance reading (0.7), the indirect instance of A-C-B might result in better results (0.9x0.9). Another application is the selection of the best

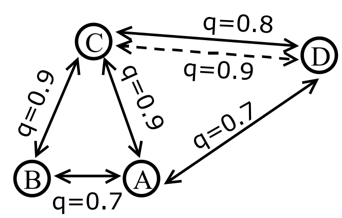

Fig. 1. Objects A-D in an environment with their quality based distance measurements

sensor readings from multiple sensors. For example, device C and D have several types of location sensors that operate in parallel: a magnetic field strength sensor and an ultrasonic sensor. Typically, a magnetic field sensor can provide high quality readings in near field, while ultrasound provides better quality at medium distances, but only when there are not reflections in the environment due to obstacles. The devices equipped with both sensors, allow the usage of a quality factor to decide which reading to process at what time. It is important to note, that often the characteristics of the sensors are not known. The quality factor that we present in this paper does not require such knowledge, but is able to determine and describe quality of a measurement outcome independent of the underlying sensor model.

A third motivating application is higher-level fusion of different types of readings. For example, from the above distance readings A-B-C and C-D, local position maps are created. The overall quality when joining them, leads to three options: Either A-B-C and C-D do not have a quality measure of their local maps, one of them has a quality measure, or both have a quality measure. Our approach is capable of handling all three situations and hence providing the best support for a fusion algorithm. It should be noted, that location map fusion can be carried out ad-hoc, without previous knowledge, and that the fusion algorithm does not need to know any characteristic of the underlying recognition method, sensor type, location or movement model, or similar in order to work correctly. In summary, the proposed approach has large potential for typical Ubicomp location systems that contain

- Devices equipped with low-tech, simple sensors
- Heterogeneous settings with various type of sensors, processing methods
- Multi-step sensor fusion
- Unknown characteristics of the sensors, processing algorithms
- Ad-hoc fusion of independently inferred location information with known execution time

For estimation of the Quality of Location (QoL) function, our system approach uses a two step self-organizing, partially supervised learning for constructing the quality measure. After learning the quality classification function - which is carried out on a PC in a training phase once for a given device or system - the classification system is able to operate directly on resource-constrained devices. It should be noted, that our system instantaneously delivers quality values for any input, spontaneously and ad-hoc. In contrast to popular Bayes-network based approaches in location system, e.g. Particle filter, our approach does not require knowledge of previous states of the overall system, e.g. a movement, to estimate the measure.

II. Related Work

Sensor fusion is widely applied for improving the location estimation. Kalman filter [1] utilizes sensor fusion of partly redundant measurements and from different sources, e.g. inertial system and GPS, to minimize the measurement errors and to reduce noise. The authors in [2] applied a distributed Kalman filter across a column of cars to improve location estimations when merging each other's local distance and a global GPS derived location. Dead Reckoning [3] predicts future positions by continuing the trace of a moving object from a known position, using last known direction and speed. These approaches always compute the most probable state respective location, even if it's poor, and return it back to the rest of the system. A further qualification, e.g. the level of accuracy, is not given.

The work in [4] estimates the accuracy in terms of meter and present it to the end user. The interpretation is left to the user and is not included for algorithmic processing to improve the location estimation. In [5] the authors present a RSSI-based iterative precision-based location algorithm for wireless ad-hoc settings, which compensates statistically for measurement errors. The error is uniformly quantified, but is tightly coupled to the concrete implementation. Our approach provides a unified measure, which is independent of concrete implementation and allows to include accuracy for further processing.

Related work using machine learning techniques to evaluate sensor performance was already done for fault detection. [6] suggests the use of Auto-Associative Neural Networks to detect faulty sensors in an aircraft engine. However, instead of detecting a binary fault state and exploiting redundancy, we focus on arbitrary accuracy.

Previous work [7] exploited a Fuzzy Inference System to estimate reliability of context classification based on the available information. This paper extends this work to cover another class of sensing systems namely relative location systems.

III. QoL Estimation Process and System

The proposed QoL function is build using a two-step estimation process: First a complete unsupervised cluster building process, where only input values - e.g. sensor values or values from other sources are fed in. Second, a supervised process, where these clusters are identified and refined using a supervised method. In this second step, both input values and ground truth are presented to the system, thus enabling it to identify the correct quality value for any input value(s) to the system. By correct we mean, that the proposed approach ensures that either the quality value can be correctly identified, or the quality value of 0 is returned. After these steps, the quality analysis model can be constructed automatically, and a quality analysis module can be generated from this model automatically. The generated quality module is independent from the concrete interpretation module (figure 2) of the system, thus the same quality module may be used for different types of interpretation modules that use the same input parameters.

Our overall system architecture is depicted in figure 2. The resulting overall system can be split into two parts: A Interpretation module, where the main estimation and interpretation of measurement are carried out, and a quality analysis module. Primary input to the system are used in both modules. Primary input could be one or several cues derived from sensor readings after feature abstraction, direct sensor readings, fused sensor information or abstracted sensor information e.g. a value describing contextual parameters. The values can also be mixed, and in most cases there are several input sources, so the primary input is often a vector of (sometimes complex) values. For the quality analysis module, additional input may be added, that mostly refer to environmental influences of the quality. These inputs are used to improve the quality measure depending on the situation the system was in. Again, such input may be sensors - e.g. sensor directly from the device - abstract context information retrieved from other devices in the environment or fused or estimated cues from sensor readings after feature extraction. The output parameters of the Interpretation module are location classes or measurements, while the quality analysis module outputs a quality measure. The measured quality is a unified value, describing the quality of the measurement in the interval between [0..1]. It is important to note, that both modules can operate completely independently, and that input parameters may only or partially come from one single device, or from the same device where the modules and/or the application is running at. This gives maximum flexibility for constructing a location system.

IV. Online Rating System

The quality analysis system holding the information about the interpretation modules error needs to meet some expenses to deliver a correct result and to still be executable on an embedded device. Most of the calculation effort is excluded in the offline system identification to keep the online system as efficient and accurate as possible. A fuzzy set theory approach

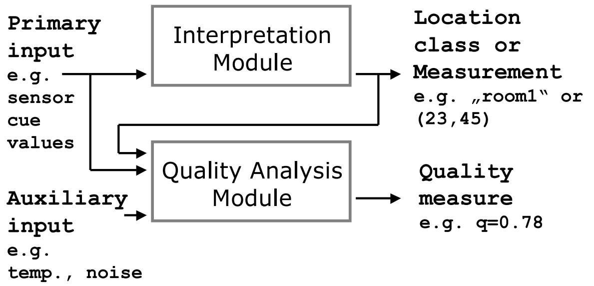

Fig. 2. Schematics of a general system for location measurement or classification with a parallel system for quality analysis.

is used for the online rating system, because the modulation is more suitable and much more flexible than probabilistic methods.

Due to various kinds of errors a measured location is only accurate to a certain degree. We reflect this fact by saying that a measurement is more or less a member of a measurement result. Note that we do not express the probability of a value. In contrast to probabilistic interpretation a membership can be expressed for several fuzzy measurements. For example, if you measure semantic location for two neighboring offices, standing between the offices is easy to express with fuzzy sets but hard and insufficiently to express with probabilistic. A fuzzy set for each office, which is needed anyway, is enough to express this situation. Between the offices a membership to both sets is non-zero. If probabilities are used there are two states representing the two offices. A situation between the offices would require a new state "between" and a probability for the state to be true - leading to a more bulky, complex and resource consuming representation. This is also one reason why fuzzy logic suits more the human understanding and thinking.

A fuzzy inference system (FIS) is a system consisting of a fuzzification, a projection of crisp values onto fuzzy sets, a set of fuzzy rules describing the relationship between input and output and a defuzzification which is a back-projection from fuzzy onto crisp. In general, each rule consists of an premise and an conclusion part. The premise is a conjunction of fuzzified values. If the conclusion consists of more than one value, the values are combined through a disjunction. The defuzzification combines all rules and projects the fuzzy outcome onto a crisp value.

A. Fuzzy Representation

A general fuzzy set $(x, \mu_{\tilde{A}}(x))$ is a tuple of a value x and a membership $\mu_{\tilde{A}}(x)$. The membership function $\mu_{\tilde{A}} : U \to [0,1]$ expresses the degree an element belongs to the fuzzy set $\tilde{A}$. In a crisp set A the membership $\mu_A : U \to \{0,1\}$ would equal 1 for all members. Typical fuzzy membership functions are gaussian-, triangular-, trapezoid-, etc. functions $\mu : U \to [0,1]$ with a maximum of one and a minimum zero. In general the fuzzy sets are an extension of the crisp set theory, and therefore fuzzy logic is an extension of the boolean logic. The fuzzy equivalent of boolean operators are functions $f : [0,1]^2 \to [0,1]$.

Applied to a fuzzy qualitity measure, which should represent the accuracy of a measurement, algorithm and/or classification, a fuzzy tuple needs to be specified accordingly as $(p_t, q_{p,t})$. The quality measure itself is expressed through the value $q_{p,t}$ of the membership function $\mu_{\tilde{p}}(\overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) = q_{p,t}$ with $\overrightarrow{\mathbf{v}}_t$ containing the information resulting in p_t, the measured location value p_t itself, and some additional information $\overrightarrow{\mathbf{u}}_t$ about environmental influences on the measurement. the entire information and the measurement are taken at time t. The question which now arises is how to obtain the quality measure $q_{p,t}$ without specifying a membership function $\mu_{\tilde{p}}$ by hand.

B. Membership Function for Quality

The quality measure is obtained through a FIS itself, which is in our case rather (ab-)used as a complex membership function $\mu(p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) \equiv \mu_{\tilde{p}}(\overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) = q_{p,t}$. It is reasonable to see the complex membership function as the combination of memberships of all possible fuzzy measurements. Looking at the complete quality management system it is this that makes it possible to integrate this FIS as a membership function into another FIS underlining the compositional features of the system. The properties of a FIS systems are advantageous for our purpose: It can be automatically constructed, it allows to easily identify unknown inputs and its rule structure guarantees a deterministic calculation effort. Automatic construction is required as it is in most cases not practical to perform manual supervised measurement runs.

Takagi, Sugeno and Kang [8][9] (TSK) fuzzy inference systems are fuzzy rule-based structures, which are especially suited for automated construction. A feature of TSK-FIS is that unknown data are mapped to the zero quality value. Note that this is in contrast to other approaches as Particle filters, where unknown data produce unforeseeable, arbitrary results. With the TSK-FIS the consequence of the implication is not a functional membership to a fuzzy set but a constant or linear function. The consequence of the rule j depends on the input of the FIS:

$$\begin{aligned} f_j(p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) \quad := \quad & \sum_{i=1}^{n} a_{ij} v_i + \sum_{i=1}^{l} a_{(i+n)j} u_i \\ & + a_{(l+n+1)j} p_t + a_{(n+l+2)j} \end{aligned}$$

The linguistic equivalent of a rule is formulated accordingly:

$$\text{IF } F_{1j}(v_1).. \text{ AND } F_{(n+1)j}(u_1).. \text{ AND } F_{(l+n+2)j}(p_t)$$
$$\text{THEN } f_j(p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t)$$

The membership functions of the rule are non-linear Gaussian functions.The antecedent part of the rule j determines the weight w_j accordingly:

$$w_j(p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) := \prod_{i=1}^{n} F_{ij}(v_i) \cdot \prod_{i=1}^{l} F_{(i+n)j}(u_i) \cdot F_{(l+n+1)j}(p_t)$$

The projection from input $\overrightarrow{\mathbf{v}}_{\tilde{p}} := (p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t)$ onto the quality measure $\widehat{q}_{p,t}$ is a weighted sum average, which is a combination of fuzzy reasoning and defuzzification. The weighted sum average is calculated according to the rules $j = 1, .., m$ as follows:

$$\mathbf{S}(\overrightarrow{\mathbf{v}}_{\tilde{p}}) := \frac{\sum_{j=1}^{m} w_j(\overrightarrow{\mathbf{v}}_{\tilde{p}}) f_j(\overrightarrow{\mathbf{v}}_{\tilde{p}})}{\sum_{j=1}^{m} w_j(\overrightarrow{\mathbf{v}}_{\tilde{p}})}$$

The TSK-FIS $\mathbf{S}$ maps onto a set $\widehat{Q}$, which is not a desirable quality measure since its boundaries cannot be determined. The value $\mathbf{S}(\overrightarrow{\mathbf{v}}_Q)$ needs to be normalized to fit in a designated set Q of quality measures, which also fits the definition of a resulting set $[0,1]$ of a membership function.

1) Normalization of FIS Result: Due to the construction of the TSK-FIS the mapping is not restricted to the particular interval $[0,1]$. The error between designated and actual output is distributed around one and zero, so values above one and below zero are possible. These values need to be normalized to the interval $Q=[0,1]$ of the desired quality measure and in conjunction to meet the demands of a membership function. The normalization is done via a function L that maps onto the interval $Q=[0,1]$ or onto an error state ε. An error state ε represents the quality measures which could not be mapped onto the interval $Q=[0,1]$ in a semantically correct way. The values lower than -0.5 would represent an error for the designated output after the normalization. A semantically correct interpretation of the value is that it belongs to zero with a mapping error. These circumstances are the same for values over 1.5 and the designated output zero. So values below -0.5 and above 1.5 are mapped with the function L onto the error state ε. Considering this the normalizing function L is defined as follows:

$$L(x) := \begin{cases} |x| & \text{if } -0,5 \leq x < 0 \\ x & \text{if } 0 \leq x \leq 1 \\ 2-x & \text{if } 1 < x \leq 1,5 \\ \varepsilon & \text{else.} \end{cases}$$

The quality measure is computed by the TSK-FIS $\mathbf{S}$ composed with the normalizing function L:

$$\mu : \begin{cases} \mathcal{V}_1 \times ... \times \mathcal{V}_N \times \mathcal{U}_L \times ... \times \mathcal{U}_L \times P & \rightarrow & Q \cup \varepsilon \\ (p_t, v_1, .., v_n, u_1, .., u_l) & \mapsto & L \circ \mathbf{S}(\overrightarrow{\mathbf{v}}_{\tilde{p}}) \end{cases}$$

The system to calculate the (normalized) quality measure is referred to as the function μ from now on. This complex membership function then gives the final measure of the location quality measure q.

V. Offline System Identification

The relationship between the system error, the input/output and environmental influences is hardly specifiable in a manual process, especially for a high diversity of influences on the system error. We therefore use a set of automatical learning and system identification algorithms to obtain the quality analyzing systems. Although it is shown that a FIS can be infinitely precise using an infinite set of rules [10], we opt for an approach that is more suitable for small sensor systems. With the chosen algorithms, a close adaptation of the system error is possible, with the resulting FIS still being applicable on small embedded sensor systems using small (8-bit) microprocessors.

A. Unsupervised Rule Identification

An unsupervised clustering algorithm is used to perform initial rule identification. Each cluster results in a fuzzy rule representing the data in the cluster and its influences on the quality analysis. Several fuzzy clustering methods are known. The methods we are looking for should - among other aspects - be able to determine the number of clusters automatically, as we do not know how many rule mappings are required. Mountain clustering [11] may be suitable, but is highly dependent on the grid structure. We opt for a subtractive clustering [12] instead. This clustering estimates every data point as a possible cluster center, so there are no prior specifications. Chiu [13] gives a description of the parameters that the subtractive clustering needs for good cluster determination. We use the subtractive clustering to determine the number of rules m, the antecedent weights w_j and the shape of the initial membership functions F_{ij}. Based on the initial membership functions a linear regression can provide the consequent functions.

B. Supervised Parameter Tuning

With a initial rule structure after the clustering, the FIS needs to be further specified. The next step is specifying the functional linear consequence of each rule for which only supervised methods are useful to converge towards an optimal mapping. In a last step of training a neural network representation of the preliminary FIS tunes the parameters to a minimum error. This tuning also requires supervision.

1) Linear Regression with Least Squares: The weights a_{ij} of the consequent functions f_j are calculated through a linear regression. The least squares method fits the functions f_i into the data set that needs to be adapted. A linear equation for the differentiated error between designated and actual output - which can be calculated with the rules and initial membership functions the subtractive clustering identified - is solved for the whole data set with a numeric method. The single value decomposition (SVD) is used to solve the over-determined linear equation. With the initial membership functions F_{ij}, the rules j and the linear consequences f_j a neural fuzzy network can be constructed. The neural fuzzy network is used to tune the parameters a_{ij}, m_{ij} and σ_{ij}^2 in an iterative training towards a minimum error.

2) Adaptive-Network-based FIS: A functionally identical representation of an FIS as a neural network is an Adaptive-Network-based FIS (ANFIS) [14]. Most of the network's neurons are operators and only the membership functions F_{ij} and the linear consequences f_j are adaptable neurons. This neural fuzzy network is used to tune the adaptable parameters a_{ij} of the linear consequences, and m_{ij} and σ_{ij}^2 of the gaussian membership functions. The tuning process is done iteratively through a hybrid learning algorithm.

3) Hybrid Learning: The learning algorithm is hybrid since it consists of a forward and a backward pass. In the backward pass we carry out a backpropagation of the error between designated and real output of the ANFIS to the layer of the Gaussian membership functions. The backpropagation uses a gradient descent method that searches a preferably global minima for the error in an error hyperplane. The forward pass performs another iteration of the least squares method with the newly adapted membership functions from the backward pass. The hybrid learning stops before an increasing of the error for a different check data set is continuously observed. The resulting ANFIS represents the qualitative non-normalized TSK-FIS $\mathbf{S}$.

VI. Application 1: Quality based Filtering

One of the most irritating features of location systems can be that the systems may suddenly jump to an entirely unrelated position and back again. To both user and consecutive processing system it can be desirable not to get such information at all. As an example, a voice guided navigation systems could easily get disturbing in such cases. Our quality measure could minimize the effect of low quality measurements thus leading to a more smooth system behavior. In this section we experimentally show how the quality measure can be used to eliminate measurement errors by filtering and how this could stabilize location systems.

A. Adaption of Discrete Systems

A discrete system recognizes locations as classes or semantic descriptions of an area, e.g. 'Office1' or 'Main Hall'. Our QoL parameter shows how much a given device - respectively its location measurement - is within that area. We carried out an experiment where RSSI measurements where used to map to a location. The system uses pPart wireless sensor nodes (Particle Computer, RFM TR1001 transceiver), RSSI measurement was based on 3 packets received by one node that are sent from another node with different sending field strength. In our experiment, sending nodes where fixed in each room sending out location packets constantly. A total of three office spaces (about 6x6 qm each) where equipped with beacon node devices.

In a first run, semantic location mapping was based on counting packets from all beacon. Classification based based on assigning the mobile nodes location to those room where the most beacon packets where counted. It is well known from literature that such assignment leads to poor results. In our test case a correct classification was observed for 45% of all cases. Although this method is simple, a filtering approach with a threshold of $\tau = 0.8$ leads to significant improvement in correct classifications ($45\% \rightarrow 92\%$).

In a second experimental run, we implemented a smarter algorithm which also takes the RSSI levels of the received nodes into account: the k-nearest-neighbors (KNN) algorithm. We measured RSSI values on anker points in the rooms and calculated representative vectors based on incoming signals from all beacon nodes. Assignment to classes - and therefore to areas in the offices - is calculated based on the similarity to the k previously defined representative vectors. The performance of this algorithm is much higher (60%) compared to packet count, but by far not useful. The improvement in error after filtering ($\tau = 0.8$) is up to 99.9%.

The best algorithm we evaluated is a FIS that finds RSSI fingerprint similarity for the different rooms. The algorithm itself correctly classifies location for 98%. Additionally a filtering can help eliminating nearly all faulty classifications. A summary of the results can be seen in figure 4. In the figure, bars indicate the percentage of correct classifications for algorithm and algorithm combined with a filter on quality measure. The quality for the three different locations for the whole test data set are shown in figure 3. The difference in quality indicates that the quality factor is state dependent, meaning some rooms are better recognizable than others.

B. Adaption of Non-Discrete Systems

The same filtering mechanisms can also be applied to non-discrete systems such as systems based on ultrasound distance measurements. In this experiment we used the BRICK devices

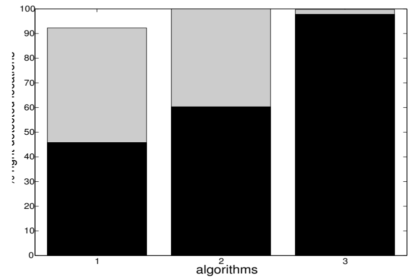

Fig. 4. Percent correctly detected semantic locations without (black) and with filtering on $q > 0.8$ (grey) for (1) packet counting, (2) K Nearest Neighbors and (3) FIS

[15] from the RELATE EU project. The device contains four narrow band ultra-sound transducers and is able to perform peer-to-peer distance and angle measure based on the time-of-flight and amplitude measurement of ultrasound pulses. In the location quality system measures a distance δ_t, a relative angle α_t and has some additional features about the environment $\overrightarrow{\mathbf{u}} = u_1, .., u_l$. The input vector for the system that performs the quality analysis is defined in the following way:

$$\overrightarrow{\mathbf{v}}_{\tilde{p}} := (p_t, \overrightarrow{\mathbf{v}}_t, \overrightarrow{\mathbf{u}}_t) = (\delta_t, \alpha_t, v_1, .., v_n, u_1, .., u_l)$$

Each time the device measures a new input $\overrightarrow{\mathbf{v}}_{t,i}$, the resulting distance, angle measurement, and environmental influences are combined with the sensor input vector into a new vector $\overrightarrow{\mathbf{v}}_{\tilde{p}}$. The quality estimation is solely based on the vector $\overrightarrow{\mathbf{v}}_{\tilde{p}}$, which is the interconnection between the location and the qualitative system. The calculation of the quality measure is done by a fuzzy inference system that holds information about the correctness of previously determined locations based on the input $\overrightarrow{\mathbf{v}}_{t,i}$.

A first training and test set 1 was collected with nodes at a constant distance of $30cm$ and an angle varying in 15° steps. The input of the quality analysis is the input the distance measure was calculated combined with the measured distance δ_t and the angle α_t. The accuracy range used is $30cm \pm 1cm$. The training of the quality measure did not show an improvement in adapting the location systems error after 3 epochs. A second quality measure was trained for an input without the angle measure α_t and evaluated with a second test set. The separation of the accurate from the less accurate distance values through the quality measure was much better for the input which includes the angle α_t.

With the density functions and the median cuts through the threshold τ the probabilities for accurate and not accurate measures can be calculated. A range is defined according to a real location value p through offset ϵ as $[p \pm \epsilon]$ [7]. For both test sets 1 and 2 the probabilities of separating the data are shown in table I. Results indicate that for a distance measure the quality always depends on the angle α_t of the receiver towards the sender.

In another test we tried to identify the quality of angle measurements. Using the same data sets as above we trained and tested the quality measure for the angle accuracy. With an accuracy range of $\alpha \pm 5^\circ$ for test set 3 (Figure 5, left plot) the accurate angles are nearly fully separable from the angles outside the accuracy range by using the quality measure and

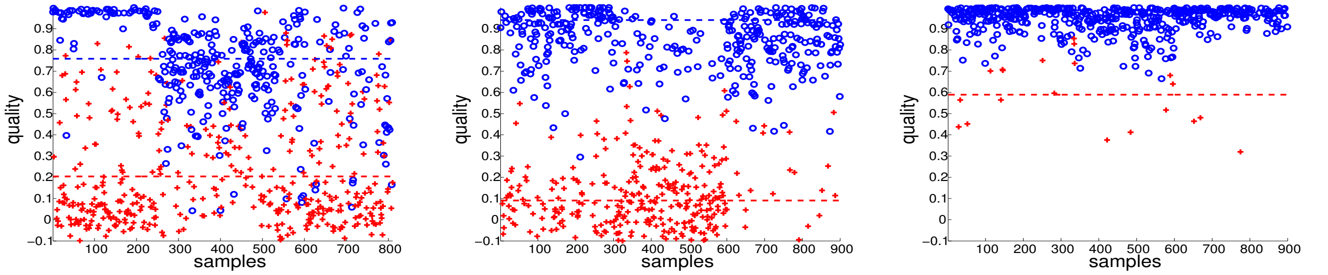

Fig. 3. Quality measure of right (o) and wrong (+) classified locations for three different algorithms: (left) packet counting, (middle) K Nearest Neighbors (KNN) and (right) Fuzzy Inference System (FIS).

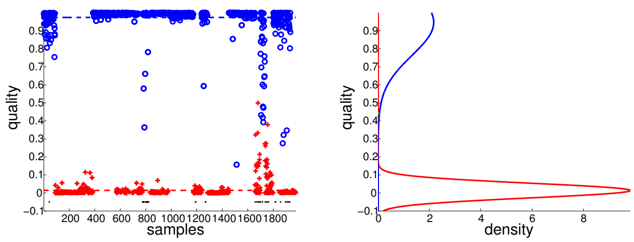

Fig. 5. Quality Measure for angle measures (left, keys like Fig. 3) for test set 3; Gaussian density functions for accurate (top) and not accurate (bottom) angle measures (right)

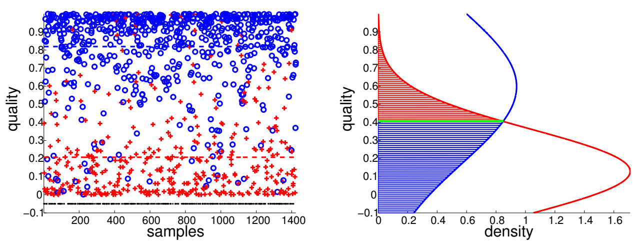

Fig. 6. Quality Measure (left, keys like Fig. 3) for test set 4; density functions (right) for accurate (top) and not accurate (bottom) measures; probabilities for false positives and negatives (shaded areas) regarding threshold τ (green line)

a threshold of $\tau = 0,194$, only a few samples are incorrectly assigned. To find the threshold τ the method described in [7] was used. The probabilities for this test set 3 are also shown in table I. Since the accuracy was previously only given for test data with a constant distance and a varying angle, a data set was collected with varying distance and angle. This leads to a test set 4. Corresponding quality measures can be seen in Figure 6. The probabilities of separating the data are also shown in table I. The results for test set 4 are not as good as the ones for set 1, but filtering based on quality still shows an improvement in the increase of accurate results. A last test 5 is used for angle α measurements and varying distances and angles. A separation of accurate and not accurate angle measurements is possible with high probabilities (table I, set 5).

C. Discussion: Classes vs. Measurements

In this section we applied a quality based filtering of semantic location classes and relative location measurements,

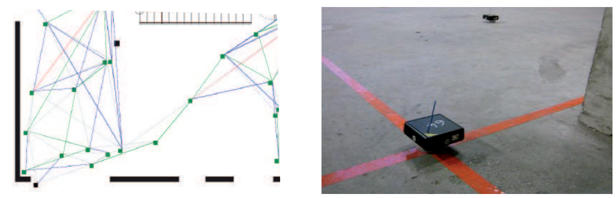

Fig. 7. Floorplan with node layout (left) and picture of node at field trial

which shows that heterogenous systems can be analyzed in quality. For semantic location classes a membership function, and therefore a quality measure is positively obtainable. The mapping of an algorithm onto a class can clearly be decided in terms of class borders and set shapes if the mapping is close to the class mean. A mapping close to the class borders is not as distinct as a centric projection. These circumstances are especially suited to be modelled with fuzzy sets. Also, the FIS can adapt this kind of system errors in a especially precise way, which is shown through the results in section VI-A.

Adapting the error of a continuous system like an ultrasound relative positioning system which was used, is in the need for a much more complex FIS than adapting classification errors. The clustering needs to identify rules covering the whole input space resulting in a increased amount of rules. Here the modality of our approach is the key of solving this problem, since the FIS can be divided into smaller parts. Many FIS combined can cover the input space and be much smaller than a big FIS covering it alone. The suitability of a FIS for getting a measurements quality is shown in section VI-B for FIS with a lot of rules or ones with a reduced set.

The statistical analysis indicates that the error characteristics of sensors and algorithms can be obtained through our quality analyzing system and be separable through it. Filtering the location measurements or classifications is on the first level of acquiring location data the only application showing the efficiency of the quality measure. The following section shows an other application, this time on the second level, for fusing location data with an improvement of the quality measure.

VII. Application 2: Location Fusion

One possible application of a quality measure is the use in graphical presentation formats. For example, a fuzzy visual representation of a point can be done using a circle [4], thus indicating the uncertainty of a position. However, the real advantages of associating a scalar quality value with

	Set 1 $\delta \wedge \alpha \rightarrow q_\delta$, $\epsilon = 1cm$ and $\tau = 0,386$	Set 2 $\delta \rightarrow q_\delta$, $\epsilon = 1cm$ and $\tau = 0,446$	Set 3 $\delta \wedge \alpha \rightarrow q_\alpha$, $\epsilon = 5°$ and $\tau = 0,194$	Set 4 $\delta \wedge \alpha \rightarrow q_\delta$, $\epsilon = 5\%$ and $\tau = 0,406$	Set 5 $\delta \wedge \alpha \rightarrow q_\alpha$, $\epsilon = 5°$ and $\tau = 0,427$
$P(\widehat{p} \in [p \pm \epsilon] \| q > \tau)$	$= 0,996$	$= 0.572$	≈ 1	$= 0,556$	$= 0,749$
$P(\widehat{p} \in [p \pm \epsilon] \| q < \tau)$	$= 0,002$	$= 0,296$	$= 2,18 \cdot 10^{-5}$	$= 0,325$	$= 0,168$
$P(\widehat{p} \notin [p \pm \epsilon] \| q < \tau)$	$= 0,996$	$= 0.572$	≈ 1	$= 0,557$	$= 0,750$
$P(\widehat{p} \notin [p \pm \epsilon] \| q > \tau)$	$= 0,001$	$= 0,132$	$= 4,39 \cdot 10^{-6}$	$= 0,118$	$= 0,082$

TABLE I
SEPARATION PROBABILITIES FOR DIFFERENT TEST SETS THROUGH QUALITY q AND THRESHOLD τ

information is its application in fusion algorithms. As depicted in figure 8 the design of our quality measure allows us to model and control the effects of fusion. In this section we present a practical application of different fusion schemes based on our quality measure. As an example of a widely spread simple fusion algorithm for location system we use the least squares method, which is typically used as linear regression in multi-lateration approaches.

In a system like the ultrasound location presented in the last section, each distance measurements δ_i can start at any position $p = (x, y)$ and measure a position (x_i, y_i) form an circle equation $(x - x_i)^2 + (y - y_i)^2 = \delta_i^2$. Having measurements of at least 3 independent locations, the position p can be described by the linear equation system $(2 * (x_i - x_0)2 * (y_i - y_0)) * p = x_i^2 - x_0^2 + y_i^2 - y_0^2 + \delta_0^2 - \delta_i^2 + \varepsilon$ of the form $Y = Xp + \varepsilon$. The least square estimation of the position $\widehat{p}$ equation can be obtained by evaluating $\widehat{p} = (X^T X)^{-1} X^T Y$.

We had previously used this algorithm to evaluate the obtainable location accuracy in a real world application trial at the training house of the Paris fire department with the BRICK hardware. Figure 7 shows the deployment of 40 nodes on two floors in a building. The nodes were deployed randomly in a 2m area of pre-recorded traces of a firemen team entering the building. The scenario was originally used to analyze the location accuracy in an ad-hoc deployed ultrasound positioning systems. The problem of such ad-hoc systems is that the sensors cannot be optimally deployed leading to a variety situation with inaccurate and fake measurements. The errors in the data set ranged from 1 to 4516 mm with an average error of 312 mm. The top curve of figure 9 shows the average and maximum errors that can be observed when using measurement from three to eleven different measurements to locate each node with the help of the others. The figure clearly shows that the fusion of more measurements can improve the error considerable by removing the systematic error.

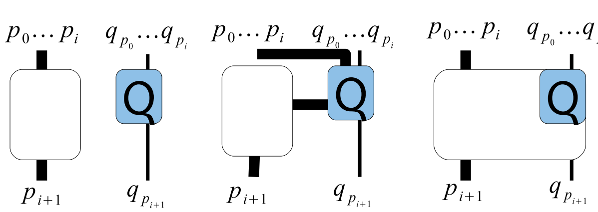

Fig. 8. Methods for integrating quality systems (Q) with a fusion algorithm: Independend(left), Filtering, Integrated (right)

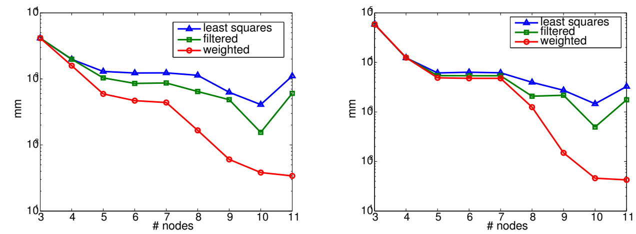

Fig. 9. Mean (left) and maximum (right) error of unweighted (Δ), weighted (+) and filtered least squares multi-lateration algorithm

A. Selective Fusion using Quality Filtering

Especially for a low number of measurements the error using least squares regression to estimate the location is still very high. We indicated as one of the problems that outliers will destabilize the system considerably. Referring back to the results from section VI we can use quality based filtering to stabilize the location algorithm. As depicted in figure 8 B, the control method remains independent of the fusion algorithm. In contrast to the previous scheme, the behavior fusion algorithm is changed by controlling its input via a threshold filter.

We took the data from a separate training data set and implemented a quality FIS for an 5% error. According to the method described in [7] we defined a filter threshold at a quality of approximately 0.5. This threshold was used to eliminate all equations with measurement of a quality below that. In case there were less than three suitable measurements we took the three best. The curves in the middle of figure 9 show the results. Except for the case of only three, where no further reduction was possible, we were able to improve the quality in all cases.

B. Quality based Fusion Control

The general problem of filtering data is that it also reduces the chance to average out an error. Other research suggest, that many inaccurate measurements can also be successfully be used to improve the error [5]. The last experiment with filtering has shown that this is not completely the case for the used regression algorithm. However, we would expect that there are cases when only unsuitable data is available using it would do no harm. In such cases it can be beneficial to use the quality measure as an input to the fusion algorithm itself. In contrast to the selection process (figure 8 A), where a quality control is implemented as an external controller, we need to adjust the algorithm and expose the quality to it.

In case of multi-lateration we can simply accommodate this fact by weighting equation by their quality. We modify the regression by changing the equation to reflect the weighted case: $\widehat{p} = (X^T W X)^{-1} X^T W Y$ and configure the weight matrix to reflect qualities of each equation. We construct the following fuzzy term $(q_{\delta_0} \vee q_{\delta_0}) \vee (q_{\delta_i} \vee q_{\delta_i})$ to describe the quality of both quadratic distances in the equation. If we use multiplication as the conjunctive fuzzy operators we get the weight matrix $W = \mathrm{diag}(q_{\delta_0}^2 * q_{\delta_i}^2)$. Figure 9 shows the results of the weighted least squares regression. As expected, we are able to combine the effects of statistic and dynamic error elimination. The weighted regression method again performs better than both previous algorithms. An interesting result is that we were able to keep the maximum error for more than 9 nodes, more than one order of magnitude below the unweighed case.

C. Discussion: Modularity vs. Integration

Integrating our quality measure into a multi-lateration scheme in various ways shows the general flexibility of the system. In this case it made sense to integrate the quality measure into the algorithm itself, which partially contradicts the modular design. However, in cases where this is not possible or complicated we can still expect improvements from filtering or at least calculate the resulting quality using fuzzy logic. Furthermore the choice of the fuzzy representation as value from 0 to 1 made it easy handle to the quality measure algorithmically.

VIII. Conclusion and Future Work

In the paper we presented a location quality measure that is generally applicable to any black-box location system. The quality system is interconnected with the actual location system only through its input and output making it possible to modularly construct and extend quality control location systems. The automated construction of the fuzzy quality system makes it adaptable to any specific characteristics of a location system and its environment. In application cases we showed that the quality measure can provide the information on the accuracy for both discrete and non-discrete location systems like RSSI-based semantic or ultrasound location at run-time, both as primary quality indicator and to control fusion.

The most definite next step will be to apply our work to more sensors and location systems. We are especially interested in finding error correlations between different sensors, so that we can take the input of, e.g. inertial sensors to estimate the quality of the ultrasound system. We hope to also be able to extract more features for a reliable classification of the error by doing some automated pre-processing and feature selection on the raw sensor data before inputting it into the quality system.

References

[1] Kalman, R.: A new approach to linear filtering and prediction problems. Journal of Basic Engineering (ASME) (1960)

[2] Hontani, H., Higuchi, Y.: mprovement of car positioning with measurements of relative distance. Fourth International Conference on Networked Sensing Systems 2007 (INSS07) (June 6-8 2007)

[3] Randell, C.; Djiallis, C.M.H.: Personal position measurement using dead reckoning. Seventh IEEE International Symposium on Wearable Computers (ISWC) (October 18-21 2003)

[4] Dearman, D., Varshavsky, A., de Lara, E., Truong, K.N.: An exploitation of location error estimation. UbiComp 2007, Springer Verlag, ISBN 3-540-74852-0 (2007) pp 181–198 Insbruck, Austria.

[5] Evers, Dulman, Havinga: A distributed precision based localization algorithm for ad-hoc networks. Pervasive Computing (2004)

[6] Mattern, D.L., Jaw, L.C., Guo, T.H., Graham, R., McCoy, W.: Using neural networks for sensor validation. AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, 34 th, Cleveland, OH (1998)

[7] Berchtold, M., Decker, C., Riedel, T., Zimmer, T., Beigl, M.: Using a context quality measure for improving smart appliances. 7th International Workshop on Smart Appliances and Wearable Computing (IWSAWC) (2007)

[8] Tagaki, T., Sugeno, M.: Fuzzy identification of systems and its application to modelling and control. IEEE Trans. Syst. Man and Cybernetics, 1985, vol SMC-15, no. 1, pp 116-132 (1985)

[9] Sugeno, M., Kang, G.: Structure identification of fuzzy model. Fuzzy Sets and Systems, 1988, vol 26(1), pp 15-33 (1988)

[10] Wang, L.X.: Adaptive Fuzzy Systems and Control. Prentice-Hall, Englewood Cliffs (1998)

[11] Yager, R., Filer, D.: Generation of fuzzy rules by mountain clustering. Journal on Intelligent Fuzzy Systems, vol 2, pp 209-219 (1994)

[12] Chiu, S.: Method and software for extracting fuzzy classification rules by subtractive clustering. IEEE Control Systems Magazine, 1996, vol. pp. 461-465 (1996)

[13] Chiu, S.: 9, Extracting Fuzzy Rules from Data for Function Approximation and Pattern Classification. In: Fuzzy Information Engineering: A Guided Tour of Applications. John Wiley&Sons (1997)

[14] Jang, J.S.R.: Anfis: Adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man and Cybernetics, 1993, vol. 23 pp. 665-685 (1993)

[15] Hazas, M., Kray, C., Gellersen, H., Agbota, H., Kortuem, G., Krohn, A.: A relative positioning system for co-located mobile devices. In: MobiSys '05: Proceedings of the 3rd international conference on Mobile systems, applications, and services, New York, NY, USA, ACM Press (2005) 177–190