Optimization-based Secure Multi-hop Localization in Wireless Ad Hoc Networks*

Sander Wozniak[1], Tobias Gerlach[2], and Guenter Schaefer[1]

1 **Telematics and Computer Networks Research Group**
Ilmenau University of Technology, Germany
sander.wozniak@tu-ilmenau.de, guenter.schaefer@tu-ilmenau.de
2 **Operations Research and Stochastics Research Group**
Ilmenau University of Technology, Germany
tobias.gerlach@tu-ilmenau.de

Abstract

The problem of localizing nodes without GPS based on a small fraction of anchor nodes which are aware of their positions is considered to be an important service for applications in wireless ad hoc networks. With an adversary trying to mislead nodes about their estimated locations, several approaches aiming to defeat attackers by means of robustness instead of cryptographic measures have been proposed in the past. Nevertheless, these robust techniques focus on single-hop based localization. Hence, we investigate the impact of employing the well-known Least Median of Squares (LMS) algorithm in the context of the multi-hop based DV-hop approach. We argue that in this case LMS is no longer able to meet its requirements. We examine the source of this behavior and show that LMS leads to more accurate results when using the median to obtain average hop lengths in DV-hop. Furthermore, we investigate the feasibility of performing lateration using the l_1-norm instead of the typically employed l_2-norm, as well as the possibility of enhancing the robustness of LMS using lateration based on the l_1-norm. Contrary to our expectations, the l_1-norm only results in a slight, neglectable advantage compared to the computationally less expensive l_2-norm lateration.

Digital Object Identifier 10.4230/OASIcs.KiVS.2011.182

1 Introduction

Wireless Ad Hoc Networks offer a wide variety of applications ranging from environmental monitoring to intrusion detection or battlefield surveillance. An important service for these applications is the localization of the participants without relying on GPS. Thus the problem of localizing nodes using only a small fraction of anchor nodes which are aware of their positions has gained much attention from researchers in the past. While these mechanisms usually assume cooperative behavior among the participants, certain applications demand the deployment of nodes in an adversarial environment. In order to prevent an adversary from misleading nodes about their locations, a variety of secure localization schemes have been presented the last few years [5]. Yet most of these approaches require single-hop communication between nodes and anchors to conduct distance measurements. In contrast, multi-hop based schemes only rely on a few anchors to measure the distance between nodes and anchors. A well-known example is the DV-hop approach, where anchors broadcast small beacon messages holding their locations [3]. Nodes receiving such a message increment a contained hop count value and broadcast the adjusted message, assuming it provides a

* This work is supported by the DFG Graduiertenkolleg 1487 *(Selbstorganisierende Mobilkommunikationssysteme für Katastrophenszenarien).*

17th GI/ITG Conference on Communication in Distributed Systems (KiVS'11).
Editors: Norbert Luttenberger, Hagen Peters; pp. 182–187

OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

shorter path to the respective anchor. Once a node has received the messages from at least three anchors, it is able to perform lateration to obtain its coordinates using the according set of references (x_i, y_i, d_i), where (x_i, y_i) is the position of the anchor and d_i the respective measured distance. In order to obtain a distance measurement d_i, the number of hops of the shortest path to the respective anchor is multiplied with an average hop length value. Anchors estimate this hop length by calculating the sum of the euclidean distances to the other known anchors and dividing it by the sum of the number of hops of the shortest paths to the according anchors. Then, the hop length is broadcast in a separate message or piggy-backed with a beacon message sent out at regular intervals. Finally, nodes receiving these estimates from several anchors calculate the mean to obtain an aggregated hop length.

Based on these observations, this work provides the following contributions: First, we investigate the influence of employing the Least Median of Squares (LMS) approach in the multi-hop based DV-hop scheme. We show that LMS is unable to defeat a basic attack in the originally described DV-hop algorithm. Second, we show that estimating an average hop length using a slightly modified technique based on the median enables LMS to again provide robustness against this attack. Finally, we investigate the feasibility of the l_1-norm in contrast to the widely-used l_2-norm, as well as the possibility of enhancing the robustness of LMS by employing the l_1-norm.

2 Linear Least Squares Lateration

Given the set of N references (x_i, y_i, d_i), a node would ideally reside at the point of intersection of at least three circles with center (x_i, y_i) and radius d_i. Hence, assuming no distance measurement errors, it would be sufficient to find this point of intersection by solving a system of non-linear circle equations [2]. In reality, however, these circles usually do not intersect at a specific location (i.e. the system of equations is not solvable). In this case, a *least squares* approach minimizing the sum of residue squares can be used to estimate a position. However, this involves solving a non-linear optimization problem, which is usually considered too expensive as it requires methods of global optimization. Therefore, the result of the non-linear least squares approach is approximated by using the *Linear Least Squares* (LLS) technique which is based on the following non-linear optimization problem [2]:

$$\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} = \arg\min_{\mathbf{x}} \|\mathbf{Ax}-\mathbf{b}\|_p \qquad \mathbf{A}\in\mathbb{R}^{(N,2)},\ \mathbf{x}\in\mathbb{R}^2,\ \mathbf{b}\in\mathbb{R}^N \tag{1}$$

Here, $\|\cdot\|_p$ is the l_p-norm ($p \geq 1$ is a parameter which may be chosen to fit a specific application) and $\mathbf{Ax} = \mathbf{b}$ is the matrix form of a system of linear equations. This system of equations is obtained by subtracting the mean of all left and right parts of the system of non-linear circle equations from each equation according to [2]. Furthermore, it is easy to see that $(\hat{x}, \hat{y})^T$ is a solution of $\min_{\mathbf{x}} \|\mathbf{Ax}-\mathbf{b}\|_p$ if and only if it is a solution of $\min_{\mathbf{x}} \|\mathbf{Ax}-\mathbf{b}\|_p^p$.

Usually, the l_2-norm ($p = 2$) is used to estimate a location. This is due to the fact that, for $p = 2$, a location can be estimated by simply solving a system of linear equations using QR-factorization for example. Nevertheless, while employing the l_2-norm ($p = 2$) is considered to be the most feasible approach, it is known to be vulnerable to malicious references forging the location or the distance to an anchor [2].

The LLS approach may also be applied to fit a function to a given set of data points. In this case, compared to the l_2-norm, the l_1-norm is generally less vulnerable to outliers contained in the data, e.g. caused by measurement errors. Therefore, we are interested in whether using the l_1-norm instead of the l_2-norm might increase the robustness of lateration

against attackers. Employing the l_1-norm, the position of a node is estimated by solving the following optimization problem:

$$\begin{pmatrix}\hat{x}\\ \hat{y}\end{pmatrix} = \arg\min_{\mathbf{x}} \|\mathbf{Ax}-\mathbf{b}\|_1 = \arg\min_{\mathbf{x}} \left\{\sum_{i=1}^{N} |\mathbf{a}_i\mathbf{x}-b_i|\right\} \tag{2}$$

This problem can be formulated as a linear optimization problem by introducing a vector $\mathbf{h} = (h_1, h_2, \ldots, h_N) \in \mathbb{R}^N$ of auxiliary variables $h_i \geq 0$, $\forall i \in \{1, \ldots, N\}$:

$$\underset{\mathbf{x}}{\text{MIN}} \left\{\sum_{i=1}^{N} |\mathbf{a}_i\mathbf{x}-b_i|\right\} \Leftrightarrow \underset{\mathbf{x},\mathbf{h}}{\text{MIN}} \left\{\sum_{i=1}^{N} h_i \;\middle|\; -\mathbf{h} \leq \mathbf{Ax}-\mathbf{b} \leq \mathbf{h}\right\}$$

Hence, we obtain:

$$\underset{\mathbf{x},\mathbf{h}}{\text{MIN}} \left\{ (\mathbf{0}^T, \mathbf{1}^T) \begin{pmatrix}\mathbf{x}\\ \mathbf{h}\end{pmatrix} \middle| \begin{pmatrix}\mathbf{A} & -\mathbf{E}\\ -\mathbf{A} & -\mathbf{E}\end{pmatrix} \begin{pmatrix}\mathbf{x}\\ \mathbf{h}\end{pmatrix} \leq \begin{pmatrix}\mathbf{b}\\ -\mathbf{b}\end{pmatrix} \right\}$$

where $\mathbf{E}$ is the identity matrix of dimension N, $\mathbf{0} = (0,0)^T \in \mathbb{R}^2$ and $\mathbf{1} = (1, 1, \ldots, 1)^T \in \mathbb{R}^N$. Depending on the norm, we refer to the respective estimation technique as l_1-LLS or l_2-LLS. While l_1-LLS is computationally more expensive than l_2-LLS, a potential increase in robustness might justify solving a simple linear optimization problem.

3 Threats and Countermeasures

There are a variety of attacks which aim at deceiving nodes about their locations [5]: *Impersonation attack*, *sybil attack*, *wormhole attack* and *location reference attack*. It should be noted here that while a Denial-of-Service attack (e.g. jamming) might also disrupt the process of localization, an adversary is usually assumed to try to unnoticeably mislead nodes about their whereabouts. Several mechanisms aiming to defeat one or more of the above mentioned threats have been proposed. They can be divided into *prevention*, *detection* and *filtering* (i.e. robust) techniques [5]. In the past, several robust location estimation schemes have been proposed. Within this work, we only consider the well-known Least Median of Squares (LMS) filtering approach [2], focusing on possible advantages of performing lateration using the l_1-norm instead of the commonly used l_2-norm, as well as its behavior in the context of the multi-hop based DV-hop scheme. In addition, we only focus on the colluding location reference attack where a number of malicious anchors broadcast beacon messages with false coordinates. These coordinates are shifted into a certain common direction away from the true positions. This threats is also known as *false beacon location attack* and very popular in terms of evaluating the robustness of location estimation schemes [2, 5].

LLS employing the l_2-norm is not robust against outliers [2]. Li et al. therefore propose to minimize the median instead of the sum of residue squares based on the method described in [4]. Finding the exact solution of this non-linear optimization problem is computationally expensive. Thus, the authors present the following algorithm as an approximate solution [2]:

1. Randomly draw $M = 20$ subsets of size 4 from the set of given references.
2. Estimate a location for each subset $j = 1, \ldots, M$ using l_2-LLS and calculate the median of the estimation residuals r_{ij}^2 to each anchor $i = 1, \ldots, N$.
3. Define $m = \arg\min_j med_i \left\{r_{ij}^2\right\}$ (least median of all medians of each subset).
4. Calculate $s_0 = 1.4826(1 + \frac{5}{N-2})\sqrt{med_i r_{im}^2}$.
5. Assign a weight w_i to each reference, where $w_i = 1$ if $|\frac{r_i}{s_0}| \leq 2.5$ or 0 otherwise.
6. Compute a weighted least squares of all given references using weights w_i. This corresponds to estimating a position using l_2-LLS with only the references with weight $w_i = 1$.

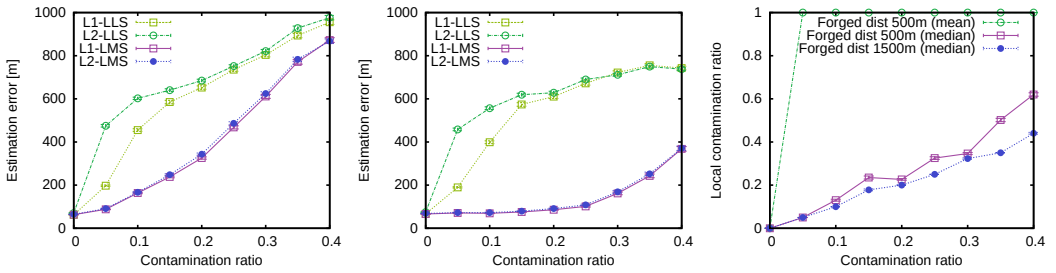
4 Evaluation

To evaluate the impact of employing LMS in DV-hop, as well as investigating the suitability of using the l_1-norm instead of the l_2-norm, we implemented DV-hop using OMNeT++ (http://www.omnetpp.org/) as follows: Each anchor broadcast a beacon message with a sequence number containing its location in a total of 3 rounds separated by intervals of roughly 60 seconds to provide the network with enough time to distribute the messages. So far, we did not incorporate a mechanism to limit flooding, since this might result in different contamination ratios at different nodes. Anchors receiving beacons from other anchors calculated a hop length estimate following either the original DV-hop approach or the median variant and included this information in the message to be sent out in the next round. Additionally, nodes kept the most recent hop length estimate announced by the corresponding anchor. Finally, after finishing the 3 rounds, nodes estimated their locations by choosing a hop length from their list of available estimates and running each of the location estimation schemes on their respective set of references. When selecting a hop length, nodes employed the mean or the median according to the technique currently in use by the anchors.

For communication among nodes, we incorporated the radio model provided by the MiXiM framework (http://mixim.sourceforge.net/). Apart from evaluating the original LMS approach using l_2-LLS in step 2 and 6 of the algorithm which from now on we will refer to as l_2-*LMS*, we also consider a new variation of LMS employing l_1-LLS which we refer to as l_1-*LMS*. We implemented and compared l_1-LLS, l_2-LLS, as well as both LMS variants. In order to obtain a location using l_1-LLS by finding a solution to (2), we employed lpsolve (http://lpsolve.sourceforge.net/). Furthermore, regarding l_2-LLS, we used lapack++ (http://lapackpp.sourceforge.net/) to estimate a position by solving the respective system of linear equations with QR-factorization. We randomly placed 300 nodes and $N = 20$ anchors on a 1000 m $\times$ 1000 m field using the uniform distribution. The field size was chosen according to a density required to prevent the partitioning of the network and to the applied transmission power of 110 mW, which roughly corresponds to an interference range of 140 m. In our scenario, the mean number of incoming connections at a node is about 12, which is a stable value above the critical threshold of 9 determined by Langendoen and Reijers [1] with a mean of about 7 hops on the shortest paths between nodes and anchors.

Regarding the false beacon location attack where anchors announce a false position, out of N anchors, $\lceil N \cdot \epsilon \rceil$ were randomly selected to be malicious. The contamination ratio ϵ was varied from 0 (no attacker) to 0.4 in steps of 5%. Malicious anchors forged their location by adding a vector defined by a common direction and specific length to their actual coordinates. The length of this vector which is from now on referred to as *forged distance* was set to 500 m and 1500 m to examine different strengths of the attack. To measure the influence of the attacking anchors, we used the *mean estimation error* which corresponds to the mean of the euclidean distance between the actual and the estimated location over all nodes. Furthermore, it should be noted here that the following figures all show the mean of 30 repetitions including the confidence intervals at a confidence level of 99 %.

We first evaluated the robustness of the location estimation schemes when employed in the original DV-hop approach. Figure 1a shows the estimation error for a forged distance of 1500 m. Here, l_2-LLS shows the expected non-robust behavior, suffering from an increasing number of malicious anchors. Furthermore, according to our assumptions, for $\epsilon < 0.2$, l_1-LLS is able to provide a decrease of the error compared to l_2-LLS. However, contrary to our initial expectations, the difference between the estimation error of l_1-LLS and l_2-LLS becomes neglectable for $\epsilon \geq 0.2$. This may be based on the fact that the system of non-linear circle



(a) Original DV-hop approach **(b)** DV-hop using median **(c)** Impact of malicious hop lengths

Figure 1 Evaluation of LMS in DV-hop (false beacon location attack).

equations is linearized by subtracting the mean of all equations according to [2]. Therefore, malicious references may still be able to influence the resulting system of equations $\mathbf{Ax} = \mathbf{b}$, preventing the l_1-norm from providing a clear advantage for an increasing forged distance.

To understand why LMS is unable to provide the expected robust behavior as shown in figure 1a, it is necessary to explain the effect of attackers exploiting honest nodes and anchors to support the attack. Malicious anchors increase the euclidean distance computed at benign anchors, while the number of hops between the anchors remains the same. Therefore, when estimating the hop length by summing up the euclidean distances and dividing it by the sum of the number of hops, an attacker is able to cause benign anchors to announce increased hop length estimates. Computed at a benign anchor, while still being influenced by malicious distance measurements, we call such a hop length estimate *polluted*. Consequently, in the original DV-hop algorithm, with the number of hops of the shortest paths being multiplied with the hop length estimate, a polluted hop length resulting from forged anchor locations affects all distance measurements obtained at a node. With the median being able to ignore outliers up to 50 %, it seems reasonable to aggregate hop length estimates using the median in order to only incorporate references from other benign anchors (which should be a majority). Zeng et al. shortly state this assumption and propose to employ the median when aggregating hop lengths at the anchors [6]. However, they do not mention the effect of benign anchors increasing the strength of the attack. Furthermore, while they provide no evaluation, we are able to confirm their assumption according to the estimation error shown in figure 1b. Here, anchors and nodes aggregate the respective hop lengths using the median, enabling LMS to yield its expected robust behavior. This decreases the estimation error from over 800 m at $\epsilon = 0.4$ in the original DV-hop to about 400 m when using the median. It should be noted here that the assumption of benign anchors supporting the attack is also confirmed by the estimation error of the non-robust l_2-LLS approach decreasing from roughly 1000 m (figure 1a) at $\epsilon = 0.4$ to about 750 m (figure 1b) when using the median in DV-hop.

Accordingly, while employing the l_1-norm reduces the estimation error for $\epsilon < 0.2$, LMS does not benefit from using l_1-LLS (figure 1a and 1b). This may be based on the fact that for $\epsilon < 0.2$, the original l_2-LMS approach is already able to filter out the majority of malicious references. Hence, in terms of LMS, we conclude that employing l_1-LLS instead of the computationally less expensive l_2-LLS approach does not provide a clear benefit.

In order to obtain a better understanding of the actual filtering ability of the median regarding polluted hop lengths resulting in an increased strength of the attack, we investigated the *local contamination ratio*. At a node, the local contamination ratio describes the ratio of malicious references among all given references. To obtain this ratio, we tagged each beacon message with a contamination field in our simulation. Hence, this field allowed to determine

whether a beacon contained a false location or a polluted hop length. While the local contamination ratio usually corresponds to the global contamination ratio ϵ, transforming a number of hops to a distance by multiplying it with a polluted hop length can result in a local contamination ratio of 1 (i.e. all references are affected by the attack). Figure 1c shows the mean local contamination ratio over all nodes. According to our expectations, when using the original DV-hop approach, all references are either malicious or affected by a polluted hop length at a forged distance of 500 m. In contrast, using the median allows to filter the majority of polluted hop lengths. For a forged distance of 500 m, the median is unable to filter all polluted hop lengths due to noise in the benign hop lengths. However, with an increasing forged distance of 1500 m, the median is able to almost filter out all polluted hop lengths. This may be based on the fact that in this case polluted hop lengths are larger than the noise among the benign hop lengths. We therefore conclude that employing the median should be considered mandatory in DV-hop.

5 Conclusions and future work

In this work, we evaluated the robustness of LMS when employed in the multi-hop based DV-hop approach. We showed that LMS is already unable to provide the expected robust behavior for a simple anchor-based attack. However, when estimating a hop length with a slightly modified approach based on the median, LMS shows the expected robust behavior. Thus, in terms of secure localization, using the median based technique in DV-hop should be considered mandatory. Furthermore, we employed the l_1-norm instead of the l_2-norm to perform lateration. Contrary to our expectations, the l_1-norm which is typically more robust against outliers, only provided a slight, neglectable benefit when employed in LMS. We assume that this behavior is based on the subtraction of the mean of all equations from each equation (i.e. the linearization of the system of equations) according to [2]. Therefore, we recommend using the computationally less expensive l_2-LLS approach.

In our future work, we aim at providing an extensive comparison of the performance of a wider variety of robust location estimation techniques employing the median-based DV-hop algorithm in a three-dimensional, multi-hop based environment.

References

1 K. Langendoen and N. Reijers. Distributed localization in wireless sensor networks: A quantitative comparison. *Computer Networks*, 43(4):499 – 518, 2003. Wireless Sensor Networks.

2 Z. Li, W. Trappe, Y. Zhang, and Badri Nath. Robust statistical methods for securing wireless localization in sensor networks. In *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 91 – 98, April 2005.

3 D. Niculescu and B. Nath. Ad hoc positioning system (APS). In *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, volume 5, pages 2926 –2931 vol.5, 2001.

4 P.J. Rousseeuw and A.M. Leroy. *Robust regression and outlier detection.* Wiley-IEEE, 2003.

5 Y. Zeng, J. Cao, J. Hong, S. Zhang, and L. Xie. Secure localization and location verification in wireless sensor networks: A survey. *The Journal of Supercomputing*, pages 1–17, 2010. 10.1007/s11227-010-0501-4.

6 Y. Zeng, S. Zhang, S. Guo, and X. Li. Secure hop-count based localization in wireless sensor networks. In *CIS '07: Proceedings of the 2007 International Conference on Computational Intelligence and Security*, pages 907–911, Washington, DC, USA, 2007. IEEE Computer Society.