

## Research Article

# A Support Vector Data Description Approach to NLOS Identification in UWB Positioning

**Shiwei Tian, Luwen Zhao, and Guangxia Li**

*College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China*

Correspondence should be addressed to Shiwei Tian; [tianxwell@163.com](mailto:tianxwell@163.com)

Received 18 February 2014; Revised 2 May 2014; Accepted 3 May 2014; Published 19 May 2014

Academic Editor: Evangelos J. Sapountzakis

Copyright © 2014 Shiwei Tian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Non-line-of-sight (NLOS) propagation is one of the most important challenges in radio positioning, and, in recent years, significant attention has been drawn to the identification and mitigation of NLOS signals. This paper focuses on the identification of NLOS signals. The authors consider the NLOS identification problem as a one-class classification problem and apply the support vector data description (SVDD), providing accurate data descriptions utilizing kernel techniques, to perform NLOS identification in ultrawide bandwidth (UWB) positioning. Our work is based on the fact that some features extracted from the received signal waveforms, such as the kurtosis, the mean excess delay spread, and the root mean square delay spread, are different between line-of-sight (LOS) and NLOS signals. Numerical simulations are performed to demonstrate the performance, using a dataset derived from a measurement campaign.

## 1. Introduction

The availability of positional information plays an important role in many applications, such as commercial, public safety, and military sectors [1]. Currently, the global positioning system (GPS) is the most important technology for providing location awareness around the globe through a constellation of at least 24 satellites [2]. Unfortunately, since the received GPS signals have extremely low power, they can be easily obstructed. For instance, GPS-based techniques fail to provide satisfactory performance due to signal blockage in many scenarios such as in urban terrain, inside buildings, and in forests or jungles. At the same time, GPS receivers are also susceptible to jamming and unintentional interference for this same reason [3].

To overcome the above problems, ground-based wireless positioning system [4], which can be served as an augmented or complementary system to GPS, has been introduced to provide positional information. Driven by the attractive characters of ultrawide bandwidth (UWB) transmission in cluttering environments, such as superior multipath resolution, obstacles-penetrating propagation, and high-resolution

ranging capabilities [5–7], UWB techniques have attracted more and more attention in the field of positioning.

However, UWB systems still face a number of technical challenges such as signal acquisition, multiuser interference, multipath effects, and non-line-of-sight (NLOS) propagation in practical scenarios [8, 9]. The latter issue, namely, NLOS propagation, which results in positive bias to the distance measurement, has been recognized as one of the major challenges for high-resolution localization systems. It is noted that the scenarios as described previously where NLOS conditions frequently occur are usually when GPS does not work well either. Hence these areas are where UWB systems are expected to work in. Therefore, it seems rather necessary to identify the NLOS signals and to mitigate their effects on positioning systems.

The goal of NLOS identification is to determine whether the range measurement between a transmitter and a receiver is biased or not. The topic has been discussed extensively in the literature, both within the cellular network and within UWB-based network. For more information about NLOS identification and mitigation, we refer the readers to [8–13] and the references therein.

It is worth noting that, in [8], a nonparametric approach relying only on UWB waveforms measurement to NLOS identification has been evaluated. The approach in [8] performs NLOS identification based on support vector machines (SVM), without requiring a statistical characterization of waveforms under LOS and NLOS conditions. In their following publication [9], the authors propose two nonparametric regression techniques to estimate the ranging error, based solely on the received waveform and the estimated distance, utilizing machine learning tools.

SVM is one of the most widely used classification techniques. However, it requires at least two classes of inputs and labels, and unfortunately, in some cases, only one class of samples can be acquired (due to the complexity or the expensive costs). Furthermore, due to diverse defect modes (NLOS due to different obstructed modes; NLOS under various environments, etc.) and their occurrence frequencies, the true distribution of the defective received signal patterns is difficult to obtain. On the contrary, normal signals are easy to collect and they involve only small variations in uniformity, as shown in Figure 1. Hence, the NLOS identification problem can be regarded as a normal received signal (LOS signal) and a defective received signal (NLOS signal) classification problem, and one-class classification is more appropriate than two-class classification for solving the NLOS identification problem.

In this paper, we apply the SVDD [14] to perform NLOS identification in ultrawide bandwidth (UWB) positioning. The SVDD is one of the best-known one-class support vector learning methods, in which one tries the strategy of using balls defined on the feature space in order to distinguish a set of normal data from all other possible abnormal objects [15]. Compared to the existing SVM-based method, our approach has the added benefit that it can be applied when only the LOS training data is obtained. To the best knowledge of the authors, this is the first time to consider NLOS identification as a one-class classification problem (also called outlier detection or novelty detection). Our finding is validated by the dataset of UWB waveforms.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation. In Section 3, the proposed techniques for NLOS identification are described. In Section 4, we perform numerical simulations based on a dataset derived from measurements. Finally, conclusions are given in Section 5.

## 2. Problem Formulation

Consider a positioning system utilizing UWB techniques, including two types of nodes: beacons are nodes with known positions, while devices are nodes with unknown positions.

Let  $M$  be the set of devices and  $S$  the set of beacons; denote by  $|M|$  the cardinality of  $M$  and by  $|S|$  the cardinality of  $S$ . Positional state of beacon  $s \in S$  and that of device  $m \in M$  are indicated, respectively, by  $\mathbf{x}_s = [x_s \ y_s \ z_s]$ ,  $s = 1, \dots, |S|$  and  $\mathbf{x}_m = [x_m \ y_m \ z_m]$ ,  $m = 1, \dots, |M|$ . A set of unknowns  $\mathbf{U} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{|M|}]$  are to be determined.

Under the assumption that the first  $L$  range measurements are biased, range measurements between device  $m$  and a beacon  $s$  can be expressed as

$$r_{sm} = \begin{cases} \|\mathbf{x}_s - \mathbf{x}_m\| + b_m + v_{sm} & s = 1, \dots, L \\ \|\mathbf{x}_s - \mathbf{x}_m\| + v_{sm} & s = L + 1, \dots, |S|, \end{cases} \quad (1)$$

where the symbol  $\|\cdot\|$  denotes Euclidean distance,  $m \in M$ ,  $s \in S$ ;  $b_{sm}$  denotes the relative NLOS measurement error between them, expressed in distance units;  $v_{sm}$  are measurements noise.

The basic idea in this paper is utilizing the received signal waveforms to identify whether the range measurement  $r_{sm}$  is biased or not, based on features extracted directly from the received waveform  $r(t)$ . The following features are considered in the existing literature: energy of the received signal (energy); maximum amplitude of the received signal (max); rise time (rise time); mean excess delay (MED); root mean square (RMS) delay spread (RMS-DS); kurtosis.

## 3. NLOS Identification Building on SVDD

Several features have been extracted directly from the received waveform, in order to capture the differences between LOS and NLOS signals. In [10], the kurtosis, mean excess delay, and RMS delay spread are utilized to perform NLOS identification. Four cases, namely, residential case (case 1), indoor office case (case 2), outdoor office case (case 3), and industrial case (case 4), are considered, and the log-normal probability density functions (PDFs) for the above four cases of the IEEE 802.15.4a channel models are demonstrated in Figure 1. The results derived in [8] demonstrate that the set of features, namely, energy of the received signal, rise time, and kurtosis, provides a good complexity/performance trade-off.

In this section, we propose a technique to identify the NLOS signals. Our technique is nonparametric and based on SVDD. We first describe the features for distinguishing LOS and NLOS situations, followed by a brief introduction to SVDD. We then describe how SVDD can be used for NLOS identification in positioning applications, without needing to determine parametric joint distributions of the features for both the LOS and NLOS conditions and without needing to obtain training dataset of NLOS as well.

*3.1. Feature Selection for NLOS Classification.* We have selected several features that have been presented in the literature. These features were selected based on the consideration about numerical simulation in Section 4. The features we will consider are as follows:

- (i) energy of the received signal:

$$\varepsilon_r = \int_{-\infty}^{+\infty} |r(t)|^2 dt, \quad (2)$$

- (ii) maximum amplitude of the received signal:

$$r_{\max} = \max_t |r(t)|, \quad (3)$$

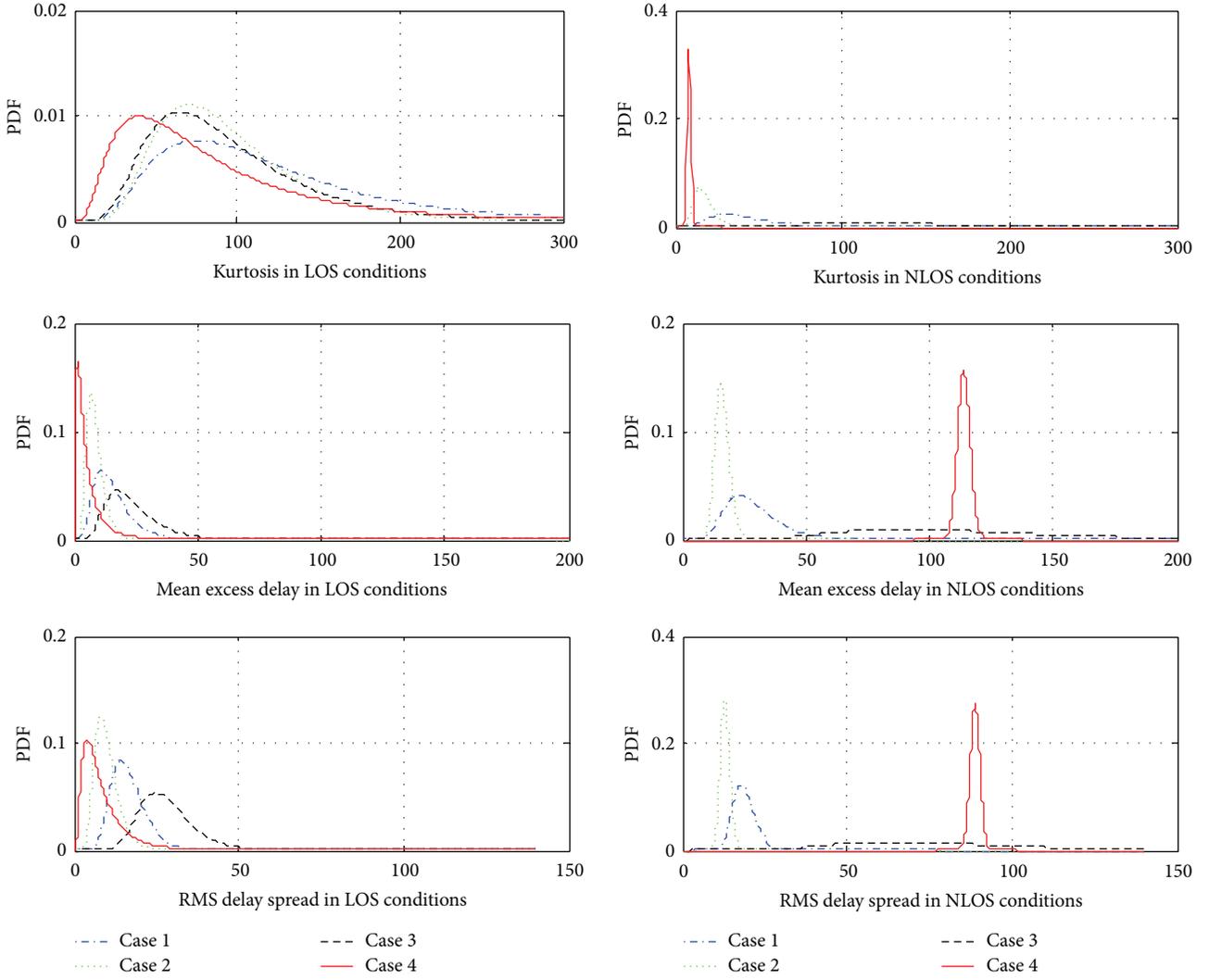


FIGURE 1: Log-normal PDFs of the kurtosis, mean excess delay, and RMS delay spread for various scenarios of the IEEE 802.15.4a channel models.

(iii) kurtosis:

$$\kappa = \frac{1}{\sigma_{|r|}^4 T} \int_T [|r(t)| - \mu_{|r|}]^4 dt, \quad (4)$$

where  $\mu_{|r|} = (1/T) \int_T |r(t)| dt$ ;  $\sigma_{|r|}^2 = (1/T) \int_T [|r(t)| - \mu_{|r|}]^2 dt$ ;  $T$  is the observation window,

(iv) mean excess delay:

$$\tau_{\text{MED}} = \int_{-\infty}^{+\infty} t \psi(t) dt, \quad (5)$$

where  $\psi(t) = |r(t)|^2 / \epsilon_r$ ,

(v) RMS delay spread:

$$\tau_{\text{RMS}} = \int_{-\infty}^{+\infty} (t - \tau_{\text{MED}})^2 \psi(t) dt, \quad (6)$$

(vi) rise time:

$$t_{\text{rise}} = t_H - t_L, \quad (7)$$

where

$$t_L = \min \{t : |r(t)| \geq \alpha \sigma_n\}, \quad (8)$$

$$t_H = \min \{t : |r(t)| \geq \beta r_{\text{max}}\},$$

and  $\sigma_n$  is the standard deviation of the thermal noise. The values of  $\alpha$  and  $\beta$  are chosen empirically in order to capture the rise time. In this paper, the values are chosen as the same as in [8].

3.2. SVDD. SVDD, which only depends on a few target objects, is one of the best-known support vector learning methods for one-class classification problems.

For a dataset containing NLOS training samples  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ , SVDD attempts to find a spherically shaped boundary, with center  $\mu$  and radius  $R$ , which contains as many LOS observations as possible while keeping as much minimum volume as possible [16]. To allow the NLOS in training data, a variable  $\xi_i$  (called slack variable) is introduced to penalize outliers for the largest distance between  $\mathbf{x}_i$  and  $\mu$ . Then, the primal formulation of an optimization problem is given as follows:

$$\begin{aligned} \min_{R, \mu, \xi_i} \quad & R^2 + C \sum_{i=1}^N \xi_i \\ \text{s.t.} \quad & \|\mathbf{x}_i - \mu\|^2 \leq R^2 + \xi_i, \quad \xi_i \geq 0, \quad \forall i, \end{aligned} \quad (9)$$

where the regularization parameter  $C$  gives the trade-off between the number of errors made on the training set and the size of the sphere description. By introducing the Lagrange multipliers for each inequality constraint of (9), the Lagrangian function is constructed as

$$\begin{aligned} L(R, \mu, \xi_i) = & R^2 + C \sum_{i=1}^N \xi_i \\ & - \sum_{i=1}^N \alpha_i \{R^2 + \xi_i - (\|\mathbf{x}_i\|^2 - 2\mu \cdot \mathbf{x}_i + \|\mu\|^2)\} \\ & - \sum_{i=1}^N \beta_i \xi_i, \end{aligned} \quad (10)$$

where Lagrange multipliers  $\alpha_i \geq 0$  and  $\beta_i \geq 0$ . Setting partial derivatives of  $L(R, \mu, \xi_i)$  to zeros, the following equations are obtained:

$$\begin{aligned} \frac{\partial L}{\partial R} = 0 : \quad & \sum_{i=1}^N \alpha_i = 1, \\ \frac{\partial L}{\partial \mu} = 0 : \quad & \mu = \frac{\sum_i \alpha_i \mathbf{x}_i}{\sum_i \alpha_i} = \sum_i \alpha_i \mathbf{x}_i, \\ \frac{\partial L}{\partial \xi_i} = 0 : \quad & C - \alpha_i - \beta_i = 0. \end{aligned} \quad (11)$$

The corresponding dual formulation is given by

$$\begin{aligned} \max_{\alpha_i} \quad & \sum_{i=1}^N \alpha_i (\mathbf{x}_i \cdot \mathbf{x}_i) - \sum_{i,j=1}^N \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C, \quad \sum_{i=1}^N \alpha_i = 1, \quad \forall i. \end{aligned} \quad (12)$$

According to the Karush-Kuhn-Tucker condition, the following equation should hold at the optimal solution:

$$\alpha_i (\|\mathbf{x}_i - \mu\|^2 - R^2 - \xi_i) = 0, \quad \forall i. \quad (13)$$

From (13), we indicate that SVDD needs only samples  $\mathbf{x}_i$  with  $\alpha_i > 0$ , namely, support vectors, which are on

the boundary or outside the hypersphere. Once SVDD is constructed on the training data, we need to decide whether a given test sample  $\mathbf{z}$  is LOS or NLOS. The criterion for the decision can be stated as

$$f(\mathbf{z}) = I(\|\mathbf{z} - \mu\|^2 \leq R^2), \quad (14)$$

where  $I$  (condition) equals one if the condition is true, otherwise zero.

In order to express more complex decision regions, kernel tricks can be used in SVDD by replacing inner products of samples  $(\mathbf{x}_i \cdot \mathbf{x}_j)$  by a kernel function  $K(\mathbf{x}_i, \mathbf{x}_j)$ . In this paper, Gaussian kernel function is utilized, which is expressed as follows:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right). \quad (15)$$

In NLOS identification problem, the receiver received several signals from the anchors at some time instant, with both LOS and NLOS signals included. Note that the objective of SVDD is to find the support of the LOS objects, and anything outside the support is viewed as NLOS. To identify the NLOS signals, we just need to obtain the training data for the LOS signals, and, after the training procedure, a decision could be made whether the received signals are the LOS signals or not.

#### 4. Numerical Simulation

In this section, to investigate how the proposed method works, we perform numerical simulations based on a dataset of UWB waveforms. For the purpose of comparison, we evaluate the detection performance of the LSSVM utilized in [8]. At the same time, example of how the classification performance is affected by the number of training samples is given. We use 10-fold cross-validation to assess the performance of the LSSVM and SVDD.

In 10-fold cross-validation for LSSVM, both LOS and NLOS signals in the dataset are utilized, while, for SVDD, only the LOS signals in the dataset are utilized and randomly partitioned into 10 parts of equal size. The SVDD is trained on 9 parts and the performance is evaluated on the remaining LOS part and a randomly selected NLOS set with an equal size.

*Data Overview.* We compare the detectors using a dataset derived from a measurement campaign, which is performed by the Wireless Communication and Network Sciences Laboratory of MIT during fall 2007. The dataset provides 1024 measurements, including 512 LOS and 512 NLOS signals. For more information about the channel model and the UWB measurement campaign, we refer the reader to [8, 9].

*Preprocessing.* There are two preprocessing steps involved. Firstly, for reasons of numerical stability, features in the dataset are converted to the log-domain before training and evaluating. Then, for the sake of averaging the results of 10-fold cross-validation, a step was taken to reduce the total

TABLE 1: The AUC values for the LSSVM and SVDD.

Subset of features	AUC value	
	LSSVM	SVDD
Set A	0.93195	0.91276
Set B	0.96186	0.94702
Set C	0.96401	0.94852
<b>Set D</b>	<b>0.96463</b>	<b>0.95233</b>
Set E	0.95894	0.95229
Set F	0.96286	0.95179

number of LOS and NLOS signals, respectively, to divide the dataset into ten equal parts.

*Performance Evaluation.* The detectors are evaluated by comparing their receiver operating characteristics (ROC) curves and the responding area-under-the-ROC curve (AUC) value. The ROC curve shows how the fraction true positive (i.e., LOS signals accepted rate in this paper, the rate deciding LOS when the signal was LOS) varies for varying fraction false positive (i.e., NLOS signals accepted rate in this paper, the rate deciding LOS when the signal was NLOS). AUC is a way to summarize a ROC curve in a single number, and higher values indicate a better separation between LOS and NLOS signals.

In Figure 2, the ROC curves are provided to evaluate the ability of the SVDD to detect NLOS signals compared to the existing LSSVM. Several subsets of the features are considered, namely, set A ( $r_{\max}$ ), set B ( $r_{\max}, t_{\text{rise}}$ ), set C ( $\epsilon_r, t_{\text{rise}}, \kappa$ ), set D ( $\epsilon_r, r_{\max}, t_{\text{rise}}, \kappa$ ), set E ( $\epsilon_r, r_{\max}, t_{\text{rise}}, \tau_{\text{MED}}, \kappa$ ), and set F ( $\epsilon_r, r_{\max}, t_{\text{rise}}, \tau_{\text{MED}}, \tau_{\text{RMS}}, \kappa$ ). The corresponding AUC values are given in Table 1.

From Table 1, it can be seen that, for both LSSVM and SVDD algorithms, set D provides the best performance among the several subsets of the features. As is shown in Figure 3, from the viewpoint of ROC curve, it can be seen that, in the region that NLOS signals accepted rate is smaller than 0.04, SVDD outperforms LSSVM, and vice versa. Utilizing the merit of AUC, we can observe that the performance of SVDD is comparable to the LSSVM, while the latter is somewhat better.

In Figure 4, the performance of SVDD as a function of the number of available LOS samples is shown. It can be observed that the case utilizing 510 LOS samples yields the best performance, and the detection performance of the SVDD detectors generally increases with the available samples size.

## 5. Conclusion

In this paper, we presented a novel NLOS identification approach building on SVDD. The technique was based on the fact that there exists some dissimilarity between LOS and NLOS signals, and several features extracted from the received waveforms can be utilized to indicate the dissimilarity. Contrary to the existing parametric approaches, the proposed technique does not rely on any statistical models

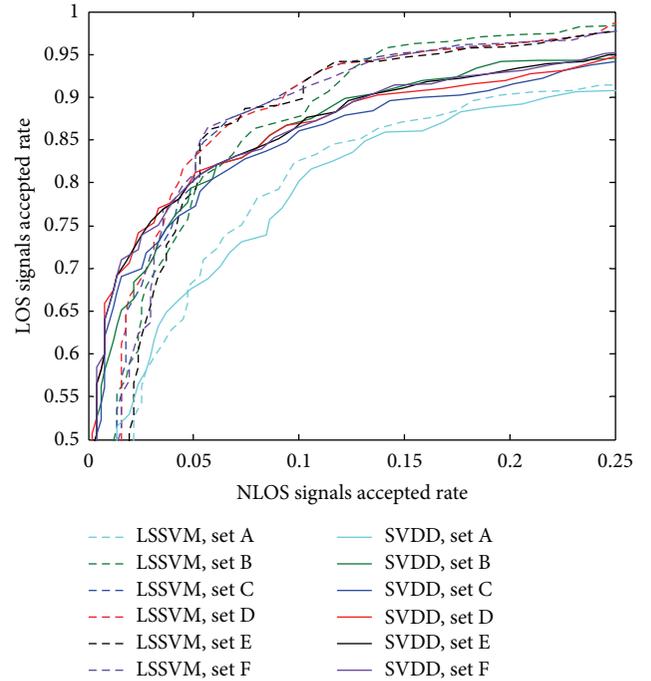


FIGURE 2: The ROC curves for the LSSVM and SVDD. For LSSVM, both LOS and NLOS signals are needed for training, and the number of each of them is 510. For SVDD, only LOS signals are needed for training, and the number of available LOS signals is 510.

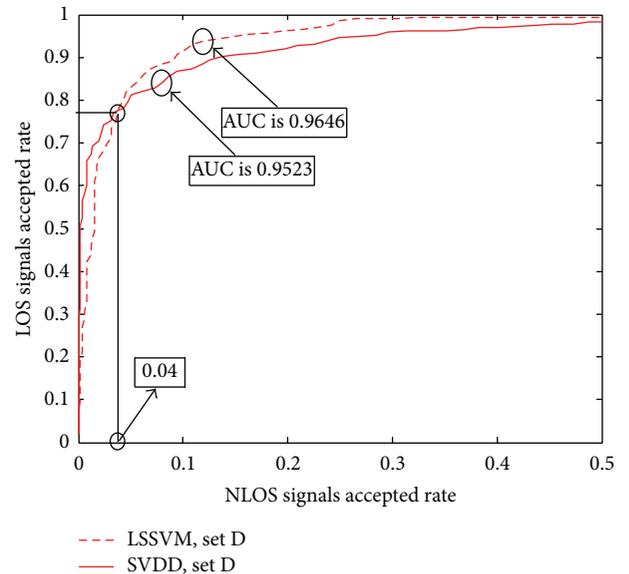


FIGURE 3: The ROC curves for the best cases in the LSSVM and SVDD algorithms.

and relies solely on received UWB waveforms. Furthermore, our technique was built on a single class compared to the existing LSSVM-based approaches, which seems quite valuable in the cases that only one class of samples can be acquired. Our results were validated by a dataset that is derived from an extensive indoor measurement campaign. As

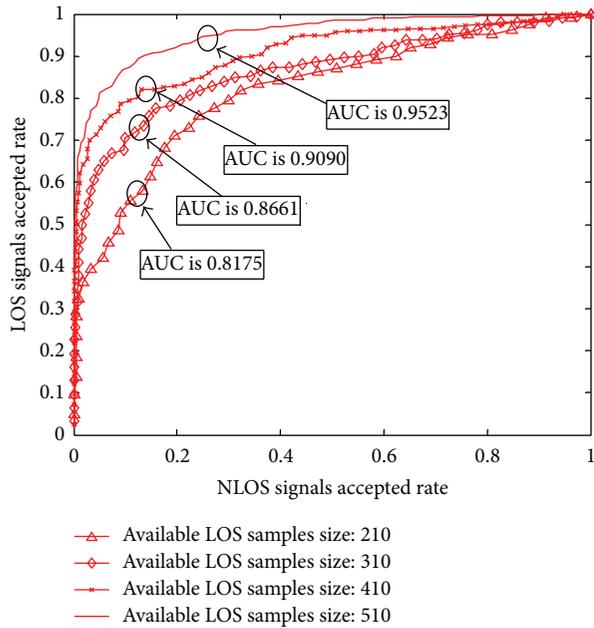


FIGURE 4: Illustration of how the performance of the SVDD detectors varies with the number of available LOS samples. Set D is utilized.

a next step, since the run-time complexity of SVDD is linear in the number of support vectors, we will develop fast SVDD to reduce the run-time complexity.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors would like to thank H. Wymeersch for providing the dataset of UWB waveforms. The authors also would like to thank D. Tax for the development and free use of the SVDD. This work is supported by the National Nature Science Foundation of China (61032004 and 91338201) and National High Technology Research and Development Program of China ("863" Program) (2012AA121605, 2012AA01A503, and 2012AA01A510).

## References

- [1] Y. Shen and M. Z. Win, "Fundamental limits of wideband localization—part I: a general framework," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4956–4980, 2010.
- [2] L. M. Kaplan, "Global node selection for localization in a distributed sensor network," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 113–135, 2006.
- [3] S. Quebe, S. DeVilbiss, J. Campbell, and C. Taylor, "Cooperative GPS navigation," in *Proceedings of the IEEE/ION Position, Location and Navigation Symposium (PLANS '10)*, pp. 834–837, May 2010.

- [4] K. G. Yu, I. Sharp, and Y. J. Guo, *Ground-Based Wireless Positioning*, John Wiley & Sons, London, UK, 2009.
- [5] S. Gezici, Z. Tian, G. B. Giannakis et al., "Localization via ultra-wideband radios: a look at positioning aspects of future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.
- [6] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 2, pp. 613–628, 2008.
- [7] T. Wang, "Cramer-Rao bound for localization with a priori knowledge on biased range measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 1, pp. 468–476, 2012.
- [8] S. Marano, W. M. Gifford, H. Wymeersch, and M. Z. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 7, pp. 1026–1035, 2010.
- [9] H. Wymeersch, S. Marano, W. M. Gifford et al., "A machine learning approach to ranging error mitigation for UWB localization," *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1719–1728, 2012.
- [10] I. Guvenc, C. C. Chong, F. Watanabe et al., "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 271984, 14 pages, 2008.
- [11] I. Güvenç and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Communications Surveys and Tutorials*, vol. 11, no. 3, pp. 107–124, 2009.
- [12] R. Casas, A. Marco, J. J. Guerrero, and J. Falcó, "Robust estimator for non-line-of-sight error mitigation in indoor localization," *Eurasip Journal on Applied Signal Processing*, vol. 2006, Article ID 43429, 2006.
- [13] M. Heidari, N. A. Alsindi, and K. Pahlavan, "UDP identification and error mitigation in ToA-based indoor localization systems using neural network architecture," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, pp. 3597–3607, 2009.
- [14] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.
- [15] J. Park, K. Daesung, J. Kim, J. T. Kwok, and I. W. Tsang, "SVDD-based pattern denoising," *Neural Computation*, vol. 19, no. 7, pp. 1919–1938, 2007.
- [16] H.-W. Cho, M. K. Jeong, and Y. Kwon, "Support vector data description for calibration monitoring of remotely located microrobotic system," *Journal of Manufacturing Systems*, vol. 25, no. 3, pp. 196–208, 2006.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

