

基于改进加权最小二乘支持向量机的 UWSN 定位

蒋 华¹, 蔡 晨², 王慧娇¹, 王 鑫^{1,2}

(1. 桂林电子科技大学 计算机信息与安全学院, 广西 桂林 541004;

2. 桂林电子科技大学 海洋工程学院, 广西 桂林 541004)

摘要: 针对水下无线传感器网络锚节点较少、迭代误差大导致的节点定位精度低的问题, 提出一种基于改进加权最小二乘支持向量机的水下三维节点定位算法; 该算法将水下三维空间分为若干立方体, 以锚节点与网格交点的距离向量作为训练集进行训练; 并利用改进的多类别模式识别方法进行分类, 以未知节点到锚节点的距离向量作为测试集确定节点坐标; 通过引入加权的思想和多类别模式识别方法增大机器学习算法的鲁棒性、降低分类次数, 从而实现水下三维节点预测定位; 仿真结果表明, 该算法在锚节点较少、网络区域较大的水下仍能保持较高的定位精度与较好的鲁棒性。

关键词: 支持向量机; 多类别模式识别; 水下无线传感器网络; 定位算法

UWSN Location Based on Improved Weighted Least Squares Support Vector Machine

JIANG Hua¹, CAI Chen², WANG Huijiao¹, WANG Xin^{1,2}

(1. College of Computer Information and Security, Guilin University of Electronic Science and Technology, Guilin 541004, China;

2. College of Ocean Engineering, Guilin University of Electronic Science and Technology, Guilin 541004, China)

Abstract: Aiming at the problem of low node positioning accuracy caused by fewer anchor nodes and large iteration errors in underwater wireless sensor networks, an underwater three-dimensional node positioning algorithm based on improved weighted least squares support vector machine is proposed. The algorithm divides the underwater three-dimensional space into several cubes, and uses the distance vector between the anchor node and the grid intersection as the training set for training. Use the classification method of multi-class pattern recognition for classification. The distance vector from the unknown node to the anchor node will be used as the test set to determine the node coordinates. By introducing weighted ideas and multi-category pattern recognition classification methods, the robustness of machine learning algorithms is increased and the number of classifications is reduced, so as to achieve underwater 3D node prediction and positioning. Simulation results show that the algorithm can maintain high positioning accuracy even in underwater areas with fewer anchor nodes and larger network areas.

Keywords: SVM; multi-class pattern recognition; UWSN; location algorithm

0 引言

由于地球资源不断地消耗, 世界各国对海洋资源日渐重视^[1]。水下无线传感器网络(UWSN, underwater wireless sensor networks)在环境监测、灾难预报、军事防御等方面都具有重要的作用, 而以上都需要建立在知道精确节点位置信息的基础之上^[2]。但水下环境的复杂性给水下无线传感器网络定位带来了许多挑战^[3]。因此水下节点定位成为了一个亟需解决的问题。

近年研究人员与科研机构提出了许多新的节点定位算法。金磊磊^[4]等基于传统水下测距定位算法, 利用最小二乘法提出适用于 UWSN 的多模态信息融合定位算法, 有效

提高节点定位精度。Ezzati^[5]等基于无线传感器网络中的接收信号强度(RSSI), 利用深度学习、极限学习机和自动编码器高级提取的特征来提高定位性能, 可以显著提升节点定位精度。吴艳玲^[6]提出一种根据邻居节点局部网络块的节点定位算法, 与机器学习领域中的降维方法相结合有效解决无线传感器网络在锚节点密度低下定位精度低的问题。朱芳^[7]等提出了基于快速支持向量机的大规模定位算法, 通过引入相似性度量来构造最小跨度, 分类速度明显提高, 且有效地解决了边界问题和覆盖孔问题。毛科技^[8]等提出一种基于支持向量机“一对一”节点定位算法, 通过引入分类思想来解决节点稀疏环境下的节点定位。这些算法均

收稿日期: 2021-01-06; 修回日期: 2021-01-26。

基金项目: 2019 广西自然科学基金面上项目(2019GXNSFAA245053); 广西科技重大专项(AA19254016)。

作者简介: 蒋 华(1963-), 男, 河南信阳人, 博士, 教授, 主要从事数据库系统及信息安全相关方向的研究。

王慧娇(1976-), 女, 辽宁铁岭人, 硕士, 副教授, 主要从事无线传感器网络及信息安全相关方向的研究。

王 鑫(1976-), 男, 陕西西安人, 硕士, 教授, 主要从事无线传感器网络及云计算技术相关方向的研究。

引用格式: 蒋 华, 蔡 晨, 王慧娇, 等. 基于改进加权最小二乘支持向量机的 UWSN 定位[J]. 计算机测量与控制, 2021, 29(8): 250-254.

在某一方面提高了在未知节点的定位精度, 但算法整体缺乏鲁棒性和对锚节点较少情况下定位精度的考虑。

针对目前已有的机器学习定位算法, 提出基于改进加权最小二乘支持向量机的水下三维节点定位算法 (WSTA, underwater 3D node location algorithm based on improved weighted least squares support vector machine)。WSTA 算法以加权最小二乘支持向量机为基础, 锚节点与水下空间立方体网格顶点之间的距离向量为训练集, 通过改进的多类别模式识别方法训练分类模型。将未知节点到锚节点的距离向量作为测试集预测节点坐标类别, 最终以立方体网格的质心作为未知节点的预测坐标。

1 系统模型

如图 1 所示, 在水下传感器网络中节点随机分布在边长为 D 的三维空间 $[0, D] \times [0, D] \times [0, D]$ 内, 黑色节点代表锚节点, 白色节点代表未知节点。将水下三维空间分为相同大小的立方体, 每一立方体代表一个类别。假设网络中 N 个节点为 $\{s_1, s_2, \dots, s_N\}$, 其中有 $\{s_1, s_2, \dots, s_k\}$ 是锚节点, 锚节点通过浮标节点和自身压力感应器预先获取自身的坐标。其余为未知坐标的普通节点 $\{s'_1, s'_2, \dots, s'_{N-k}\}$ 。节点的通信半径为 R , 且所有节点的计算能力相同。

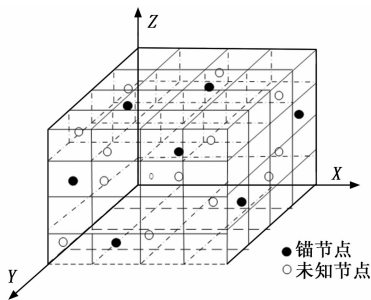


图 1 水下三维空间分类图

2 算法实现

2.1 加权最小二乘支持向量机

支持向量机^[9] (SVM, support vector machine) 是由统计学习理论发展而来, 可用于解决有效样本学习问题, 受到了越来越广泛的关注。加权最小二乘支持向量机^[10] 定位算法是由 SVM 衍生的一种无线传感器网络的定位算法。这种算法求解是线性方程的求解, 计算量小, 通讯开销较小可以适用于水下无线传感器网络节点定位。

将原 SVM 的不等式约束问题转化为等式约束:

$$\min_{w,b,e} J_p(w,e) = \frac{1}{2}w^T w + \frac{1}{2}\gamma \sum_{k=1}^N e_k^2$$

$$s. t. v_k = w^T \phi(x_k) + b + e_k, k = 1, \dots, N \quad (1)$$

其中: $\phi(x_k)$ 为非线性映射函数; b 为偏差项; w 为权重; e_k 为随机误差; γ 为正则化参数。

由于 LS-SVM 以二次损失函数作为经验风险, 这会导致模型对噪声特别敏感。因此将误差变量 $e_k = a_k/\gamma$ 引入权值 μ_k 。

构建拉格朗日函数如 (2) 所示:

$$L(w,b,e,a) = 2^{-1}w + 2^{-1}\gamma \sum_{k=1}^N \mu_k e_k^2 - \sum_{k=1}^N a_k \{w^T \phi(x_k) + b + e_k - v_k\} \quad (2)$$

根据 KKT 条件, 将 $L(w,b,e,a)$ 分别对 w,b,e,a 求偏导, 由此:

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{k=1}^N a_k \phi(x_k) \\ \frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{k=1}^N a_k = 0 \\ \frac{\partial L}{\partial a_k} = 0 \Rightarrow a_k = \gamma \mu_k e_k \\ \frac{\partial L}{\partial e_k} = 0 \Rightarrow w^T \phi(x_k) + b + e_k - v_k = 0 \end{cases} \quad (3)$$

将式 (3) 代入式 (2) 中, 消去 w 和 e_k , 可得矩阵等式 (4):

$$\begin{bmatrix} 0 & \mathbf{1}_N^T \\ \mathbf{1}_N & \Omega + \mathbf{V}_\gamma \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ v \end{bmatrix} \quad (4)$$

其中: \mathbf{V}_γ 为对角矩阵:

$$\mathbf{V}_\gamma = \text{diag} \left\{ \frac{1}{\gamma \mu_1}, \frac{1}{\gamma \mu_2}, \dots, \frac{1}{\gamma \mu_N} \right\} \quad (5)$$

权值 μ_k 是根据误差变量 $e_k = a_k/\gamma$ 来确定。

$$\mu_k = \begin{cases} 1, & \left| \frac{e_k}{\hat{s}} \right| \leq c_1 \\ c_2 - \frac{\left| \frac{e_k}{\hat{s}} \right|}{c_2 - c_1}, & c_1 \leq \left| \frac{e_k}{\hat{s}} \right| \leq c_2 \\ 10^{-4}, & \text{otherwise} \end{cases} \quad (6)$$

常数 c_1 和 c_2 通常设定为 $c_1 = 2.5, c_2 = 3, \hat{s}$ 是误差变量 e_k 的标准方差的鲁棒估计值:

$$\hat{s} = \frac{IQR}{2 \times 0.6745} \quad (7)$$

其中: IQR 是 e_k 的 25% 与 75% 分点之间的间距, 依照 Mercer 定理, 有映射 ϕ 及核函数:

$$K(x_k, x_n) = \phi(x_k)^T \phi(x_n) \quad (8)$$

得出相应的判别函数为:

$$y(x) = \text{sgn} \left[\sum_{k=1}^N a_k K(x, x_k) + b \right] \quad (9)$$

其中: a 与 b 是通过式 (4) 计算所得。

由此可见, 权值的引入增加了算法的鲁棒性与稀疏性, 且计算量并没有增加, 十分适合水下节点定位。

2.2 基于改进加权最小二乘支持向量机算法模型

2.2.1 训练集准备阶段

假定锚节点 $S_i (i = 1, 2, \dots, k)$ 到所有划分立方体网格顶点 $K_j (j = 1, 2, \dots, l)$ 的距离设定为 $h_{ij}(S_i, K_j)$, 则锚节点 S_i 到其他所有立方体网格顶点的距离可以组成距离向量:

$$d_i[h_{i1}(S_i, K_1), h_{i2}(S_i, K_2), \dots, h_{il}(S_i, K_j)]$$

对于水下三维环境节点坐标求解, 需要对 3 个维度上的坐标进行分别求取。因此以水下环境为基础建立坐标轴, 将

X、Y、Z方向上的三维空间等分为M个类别，因此X轴上存在M个分类 $\{cx_0, cx_1, \dots, cx_{M-1}\}$ 。假定每一个分类中包含该分类上的所有节点坐标。同理，Y、Z上同样存在M个类别 $\{cy_0, cy_1, \dots, cy_{M-1}\}$ 、 $\{cz_0, cz_1, \dots, cz_{M-1}\}$ 。此时水下无线传感器网络已经被我们分解成若干个立方体，每一个立方体便代表着算法中的一个类别，锚节点的类别记为 $[cx_q, cy_l, cz_v]$ 。

锚节点的距离向量 d_i 与坐标类别构成训练集：

$$T_x = \{(d_i, cx_q) \mid i = 1, 2, \dots, k\}$$

$$T_y = \{(d_i, cy_l) \mid i = 1, 2, \dots, k\}$$

$$T_z = \{(d_i, cz_v) \mid i = 1, 2, \dots, k\}$$

2.2.2 核函数选取

支持向量机算法均需要合适的核函数。本文采用目前应用较多的高斯核函数，又名径向基（RBF, radial basis function）函数作为核函数：

$$K(x, x_k) = \exp\left(-\frac{\|x - x_k\|^2}{2\sigma^2}\right) \quad (10)$$

RBF核函数可用于线性不可分的情况。可将样本映射到更高维的空间，且具有较宽的收敛域以及唯一最佳逼近的特征。因此，RBF函数针对不同维度与大小的样本均可实现较好的分类效果。

2.2.3 训练阶段

由公式(9)可知判别函数 $y(x)$ 可用于判断 m_1, m_2 两类样本，若 $y(x) > 0$ ，则判 x 属于 m_1 类，记 m_1 类一票，否则记录 m_2 为一票，最后根据样本得票多的一方作为分类结果。但由于节点分类是一个多分类问题，对于多分类的支持向量机，则需要多个分类器。传统的“一对一”分类方式需要每两个类别的数据训练出一个二分类器，这样区分M个类别需要的分类器个数为：

$$P = M(M-1)/2 \quad (11)$$

为降低运算量，对锚节点训练集进行判别后，这里使用改进的多类别模式识别方法^[11]进行分类。通过这种方法也对类别进行多轮的筛选比较，但相比“一对一”分类方法，该方法每一轮并非将所有类别都进行比较。而是进行单循环相邻比较，每轮中均去掉得票最低的类别，留下的类别才能参加下轮比较。

如图2(a)，相邻的类别进行比较。设C为得票最少的类别，如图2(b)，在第二轮中将B与D进行比较训练，其余类别则继续沿用上一轮的结果。在各个类别两轮票数相加后，假设B为第二轮得票最少的类别。如图2(c)所示，去掉B后将A与D作为比较，其他比较仍然沿用原有的结果。

若第三轮本轮累计统计票数后D被淘汰，如图2(d)，便无需再进行新的比较。只需将之前A与E比较重得票数高的类别作为分类结果。

由此可知，若单个坐标轴方向上有M个类别，除去第一轮需要M次比较，以后每轮仅需一次比较。且由于最后一轮可根据前几轮投票结果进行判定，最终使用多类别模式识别的分类方法判断未知节点H的类别需要比较的次数为：

$$P = 2M - 3 \quad (12)$$

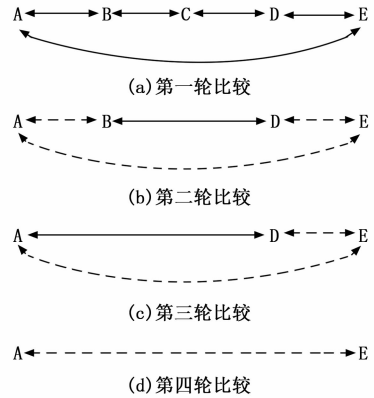


图2 改进的多类别模式识别方法的分类流程

由此可见，M的值越大，相比“一对一”投票的分类过程则会愈加简便。且改进的多类别模式识别的分类方法在不降低精度的情况下减少分类次数，其结果可靠性更好，十分适用水下无线传感器网络的节点坐标分类。

2.2.4 测试阶段

根据2.2.3节，所有锚节点训练样本分类完成且对水下网格类别进行编号，分类模型即构造成功。

设未知节点 S'_u ($u = 1, 2, \dots, N-k$) 与锚节点之间的距离为 L_i (S'_u, S_i) ($i = 1, 2, \dots, k$)，获取未知节点 S'_u 与所有锚节点之间的距离向量 d'_u ，并将其作为测试集：

$$d'_u [L_1(S'_u, S_1), L_2(S'_u, S_2), \dots, L_k(S'_u, S_k)]$$

根据训练的模型，将测试集带入进行预测未知节点的类别编号，并将该类别的质心作为未知节点的坐标，完成定位。

3 基于改进加权最小二乘支持向量机定位

根据多类别模式识别的分类方法对未知节点进行分类，判断未定位节点的分类信息，从而计算未知节点的坐标信息。基于改进加权最小二乘支持向量机的水下三维节点定位算法（WSTA算法）的核心定位流程如下：

3.1 节点定位流程

Step1: 将水下立方体环境分割为等大立方体网格，通过RSSI测距算法在每个锚节点内生成到所有立方体顶点的距离向量，生成训练集。

Step2: 在水下环境上建立三维坐标轴，并对坐标轴上的每一格进行分类。以X轴为例，设该维度上存在M个分类。其中每一分类包含X轴上，坐标在 $[q \times D/M, (q+1) \times D/M]$ 的所有锚节点，D为区域边长。同理Y轴与Z轴上的单一分类均包含Y、Z坐标在该分类的所有节点。记节点的分类类别为 $[cx_q, cy_l, cz_v]$ 。

Step3: 分别在3个维度上利用锚节点的距离向量作为训练集来训练分类器。并使用改进的多类别模式识别的分类方法进行分类。

Step4: 根据训练得到的分类器，将未知节点与所有锚节点之间的距离向量作为测试集预测未知节点类别。

Step5: 可得未知节点类别 $[cx_q, cy_l, cz_v]$ ，因此即可判断

未知节点的坐标范围: $[q \times D/M, (q+1) \times D/M] \times [t \times D/M, (t+1) \times D/M] \times [v \times D/M, (v+1) \times D/M]$ ($q, t, v \in [0, M-1]$)。计算该网格质心坐标, 将坐标作为未知节点坐标: $((q+2^{-1}) \times D/M, (t+2^{-1}) \times D/M, (v+2^{-1}) \times D/M)$ 。

3.2 定位流程

WSTA 算法定位流程如图 3 所示。

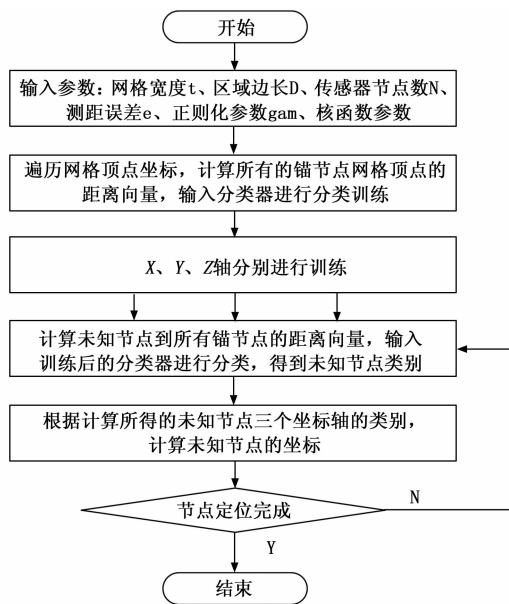


图 3 WSTA 算法定位流程图

3.3 节点定位误差

经过上述步骤进行计算后, 可预测出未知节点 H 的坐标范围。如图 4, 在该网格中, 若定位算法能正确判断未知节点 H 的所在的网格区域, 即分类正确的节点的最大误差为:

$$e = \frac{\sqrt{3}}{2} D/M$$

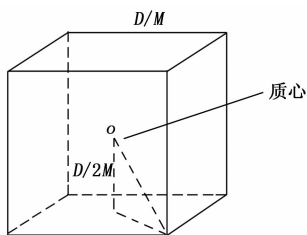


图 4 WSTA 定位误差分析图

4 实验设置与仿真

实验仿真选择在 Windows 10 PC 机上通过 MATLAB 工具实现基于改进加权最小二乘支持向量机的水下无线传感器网络定位方法。对本文的 WSTA 算法与基于 SVM 的定位算法^[7]和经典 RSSI 定位算法的平均误差率进行比较, 以及划分立方体网格宽度 t 对 WSTA 算法定位误差的影响。

仿真区域为边长分别为 80 m、100 m、120 m、140 m、160 m、180 m 的正方形区域, 随机部署 50 个传感器节点, 节点通信距离 R 为 30 m, 网络通信模型为 Regular Model,

实验参数如表 1 所示。

表 1 实验参数设置

| 参数 | 值 |
|-----------|----------------------------|
| 仿真区域边长/m | 80,100,120,140,160,180 |
| 传感器节点总数 N | 50 |
| 节点通信半径/m | 30 |
| 网格宽度/m | 8,9,10,11,12,13 |
| 正则化参数 gam | 60 |
| RBF 核函数参数 | 1 |
| 平均测距误差 | 0.05,0.1,0.15,0.2,0.25,0.3 |

设节点预测坐标为 (x', y', z') , 节点的实际坐标为 (x, y, z) 。因此节点定位误差可使用如下公式进行计算:

$$m = \frac{\sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2}}{10} \quad (13)$$

4.1 网格宽度对算法平均定位误差的影响

网格宽度是影响机器学习定位算法的一大因素。图 4 展示了在 90 M 的网络范围内 WSTA 算法和 SVM 定位算法在不同网格宽度下的平均定位误差的比较。由仿真实验结果可知, 在传感器稀薄的水下无线传感器网络中, 随着网格宽度的逐渐增大, WSTA 算法和 SVM 定位算法的平均定位误差均逐渐增大。但在 $t=12$ M 时, SVM 定位算法的平均定位误差已经将近 30%, 增长急速。WSTA 算法则较为稳定, 且整体定位精度优于 SVM 定位算法。

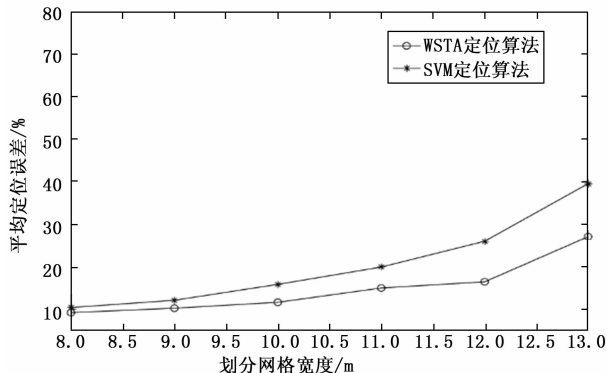


图 5 划分网格宽度对定位误差的影响

4.2 不同定位算法的比较

4.2.1 网络范围对平均定位误差的影响

图 5 是 RSSI 定位算法、SVM 定位算法、WSTA 算法 3 种定位算法的平均定位误差受网络区域边长的影响。在范围广、节点稀疏的水下无线传感器网络中, RSSI 定位算法在 80~120 m 的范围内较为平缓, 但随着区域边长变大, 定位误差急剧增加。SVM 定位算法定位误差增长较大, 对大范围环境适应性较差。WSTA 定位算法变化较为平缓, 对于范围较大的区域也能保持较好且稳定的定位精度, 鲁棒性较好。

4.2.2 网络范围对平均定位误差的影响

图 6 是 RSSI 定位算法、SVM 定位算法、WSTA 算法

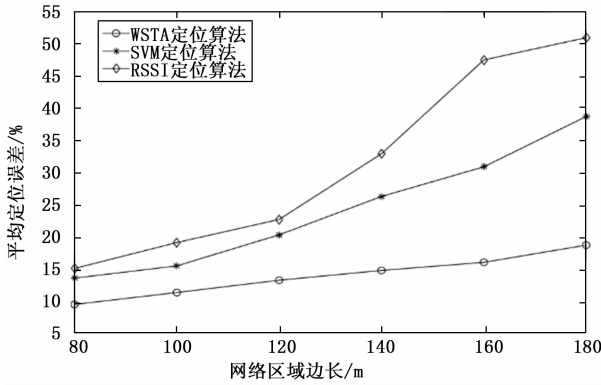


图 6 网络区域边长对定位误差的影响

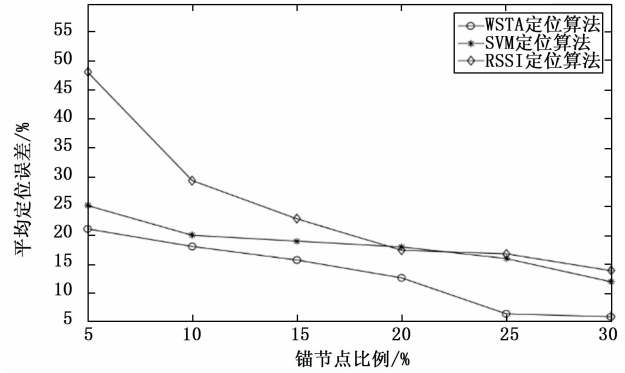


图 8 锚节点比例对定位误差的影响

是 3 种定位算法的在不同测距误差下的定位误差的变化情况。在传感器节点稀疏的水下无线传感器网络中，WSTA 算法的节点定位率最高，且节点定位率稳定，优于其他两种定位算法。在测距误差为 0.05 时，SVM 定位算法与 WSTA 算法较为接近，但随着测距误差的增加，平均定位误差急剧增加到最高。WSTA 算法在测距误差 0.2 时平均定位误差逐渐趋于平稳，受误差影响较小。

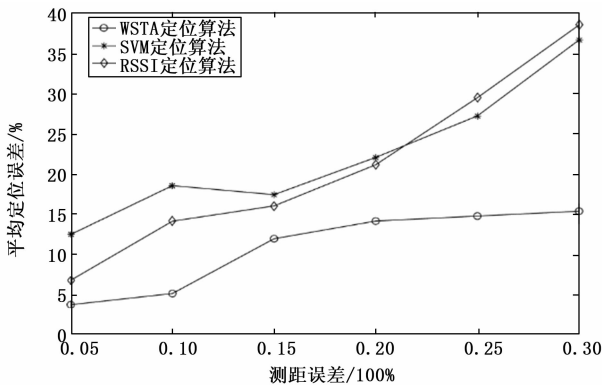


图 7 测距误差对定位误差的影响

4.2.3 锚节点比例对平均定位误差的影响

图 7 是 RSSI 定位算法、SVM 定位算法、WSTA 算法 3 种定位算法的在不同锚节点比例时定位误差的对比情况。在传感器节点稀疏的水下无线传感器网络中，RSSI 定位算法在锚节点比例 5% 时，定位误差最高。随着锚节点比例的增加，定位精度急剧提高，逐渐接近于 SVM 定位算法。WSTA 算法在锚节点比例较低时，与 SVM 定位算法较为接近。在锚节点增长到 25% 时，定位误差趋于稳定，整体精度高于其他两种算法。

在传感器节点稀疏的水下无线传感器网络中，WSTA 算法整体平均定位误差低于 SVM 定位算法与 RSSI 定位算法。环境变化对 WSTA 算法精度影响较低，算法鲁棒性高。因此，WSTA 算法整体具有较好的定位性能。

5 结束语

本文基于改进加权最小支持向量机提出一种水下三维定位算法。在该算法中，引入加权最小二乘支持向量机的

同时通过改进的多类别模式识别方法对节点进行快速分类，提高定位精度与算法鲁棒性并降低区域大小对定位精度的影响。将 WATA 算法与 SVM 定位算法、RSSI 定位算法进行仿真对比，结果证明 WATA 算法在水下传感器节点位中的确实具有较好的鲁棒性，能够有效提升水下节点定位的准确性。

参考文献:

- [1] QIAO G, ZHAO C, ZHOU F, et al. Distributed Localization Based on Signal Propagation Loss for Underwater Sensor Networks [J]. IEEE Access, 2019, 7: 112985 - 112995.
- [2] LIU J, WANG Z, CUI J, et al. A joint time synchronization and localization design for mobile underwater sensor networks [J]. IEEE Transactions on Mobile Computing, 2016, 15 (3): 530 - 543.
- [3] XING W, ZHAO Y, et al. Convergence analysis on multi-AUV systems with leader-follower architecture [J]. IEEE Access, 2017, 5: 853 - 868.
- [4] 金磊磊, 梁红, 马艳. 基于水下无线传感器阵列网络多模态信息融合的目标定位 [J]. 西北工业大学学报, 2017, 35 (6): 1020 - 1025.
- [5] CHENJI H, STOLERU R. Toward accurate mobile Sensor network localization in noisy environments [J]. IEEE Transactions on Mobile Computing, 2013, 12 (6): 1094 - 1106.
- [6] 吴艳玲. 一种基于局部信息最小二乘法的节点定位算法 [J]. 吉林大学学报 (理学版), 2017, 55 (4): 952 - 956.
- [7] ZHU F, WEI J. Localization algorithm for large scale wireless sensor networks based on fast-SVM [J]. Wireless Personal Communications, 2016, 95 (3): 1859 - 1875.
- [8] 毛科技, 范聪玲, 叶飞, 等. 基于支持向量机的无线传感器网络节点定位算法 [J]. 计算机研究与发展, 2014, 51 (11): 2427 - 2436.
- [9] CORINNA C, VLADIMIR V. Support-vector networks [J]. Mach Learn, 1995, 20 (3): 273 - 297.
- [10] 顾燕萍, 赵文杰, 吴占松. 最小二乘支持向量机鲁棒回归算法研究 [J]. 清华大学学报 (自然科学版), 2015, 55 (4): 396 - 402.
- [11] 柳长源, 毕晓君, 韦琦. 基于向量机学习算法的多模式分类器的研究及改进 [J]. 电机与控制学报, 2013, 17 (1): 114 - 118.