第 35 卷 第 6 期 计 算 机 工 程 2009 年 3 月
Vol.35 No.6 Computer Engineering March 2009
·网络与通信· 文章编号：1000—3428(2009)06—0133—03 文献标识码：A 中图分类号：TP393

改进的无线传感器网络定位算法

张 佳[1]，罗军勇[1]，王 艳[2]，姚 刚[1]
(1. 解放军信息工程大学信息工程学院，郑州 450002；2. 河南科技学院，新乡 453003)

摘 要：定位是无线传感器网络的基础工作。现有定位算法利用参考节点的位置信息对非参考节点进行定位，当该信息受到攻击或误差的影响时，将导致算法精确度下降。该文将传统最小二乘定位算法与 Metropolis-Hasting 抽样算法有机结合，提出一种改进的最小二乘定位算法。建造一个可能遭受攻击的模拟环境，在该环境下比较改进后的算法和原算法，结果表明，改进后的算法具有较好鲁棒性。

关键词：无线传感器网络；Metropolis-Hasting 抽样算法；分布特性；鲁棒性

Improved Localization Algorithm for Wireless Sensor Network

ZHANG Jia[1], LUO Jun-yong[1], WANG Yan[2], YAO Gang[1]
(1. Institute of Information Engineering, PLA Information Engineering University, Zhengzhou 450002;
2. Henan Institute of Science and Technology, Xinxiang 453003)

【Abstract】 Location is foundation work in Wireless Sensor Network(WSN). Existing location algorithms need the position information of beacon nodes to locate the non-beacon nodes. When the position information is affected by the attacker or errors, the accuracy of these algorithms will degrade. This paper combines the LS localization algorithm and Metropolis-Hasting sample algorithm, and proposes an improved LS localization algorithm. It constructs a simulation circumstance which might be attacked and compares the improved algorithm with original algorithm. Experimental results demonstrate that the improved algorithm has better robustness than original algorithm.

【Key words】 Wireless Sensor Network(WSN); Metropolis-Hasting sample algorithm; distribution trait; robustness

1 概述

无线传感器网络需要捕获客观现实的时空特性，因此，必须获取相关位置信息。拥有位置信息的无线传感器网络(Wireless Sensor Network, WSN)可以用于跟踪物体、提供向导等工作。传统定位方法，如 GPS 和人工配制方法不适应体积小、能耗低、规模大的传感器网络。因此，多数研究者假设一个网络中只有若干传感器具有GPS接收器或者可以由人工配制位置信息，此类特殊传感器称为参考节点，其他没有位置信息的节点称为非参考节点。此时，定位问题即非参考节点如何根据参考节点的位置信息进行自我定位。

目前的定位算法主要有基于测距的定位算法和非基于测距的定位算法。基于测距的定位算法需要根据所测得参考节点之间的距离或角度进行定位，一般使用最小二乘(LS)原理和三角测量法求得非参考节点的最佳估计位置，例如 TDoA, AoA 和 RIPS。非基于测距的算法无须测量距离，它根据自身与参考节点之间的相对位置进行估计，例如质心法、APIT[1], DV-HOP, Amorphous, LSBA[2]。上述 2 类定位算法需要根据参考节点的位置信息进行定位，因此，参考节点是整个传感器网络中最脆弱的地方，如果参考节点的位置信息被攻击或被干扰，那么非参考节点的估测位置将受到影响。若所有参考节点被攻击，则整个网络无法得到正确定位。本文认为对于一个传感器网络而言，某些参考节点被攻击后位置信息发生错误造成的影响远大于某些参考节点不能工作带来的影响。其原因是前者会降低非参考节点的精确度。

鉴于此，本文提出将 Metropolis-Hasting 抽样算法[3]与定位算法相结合的安全定位策略，根据受攻击后节点位置信息的统计特性和参考节点位置信息分布特性，剔除被攻击参考节点的位置信息，并保留足够的有用位置信息进行定位。

2 相关工作

近年来，无线传感器网络安全问题被越来越多地关注，攻击目标通常是攻击信号的传播测量时间、强度和角度以及攻击路由表。文献[4]给出一种验证方法防止攻击的发生。SecRLoc 对硬件进行特殊设计，利用 2 个可以连通的节点在距离上的限制条件进行抵御攻击，文献[5]从新的角度探讨抵御攻击的方法。

因为攻击手段千差万别，所以不可能设计一种可以抵御所有攻击的方法。因此，本文重点研究如何在攻击发生后，准确地提取未受感染的信息。

3 结合 Metropolis-Hasting 抽样原理的定位算法

3.1 Metropolis-Hasting 抽样算法的启发

3.1.1 马尔可夫链蒙特卡罗方法

定义 为了模拟服从给定分布的随机变量，生成一个易于实现的不可约遍历链 $X=\{X_n,n\geq 0\}$ 作为随机样本，使其平衡分布为 π 的方法，称为马尔可夫链蒙特卡罗(MCMC)方法。

3.1.2 Metropolis-Hasting 抽样算法

目前有很多 MCMC 方法，本文选择 Metropolis-Hasting 抽样算法，其基本思路如下：

任意选择一个不可约的转移概率 $q(\cdot,\cdot)$ 和一个函数

作者简介：张 佳(1979－)，男，硕士研究生，主研方向：网络安全；罗军勇，教授；王 艳，学士；姚 刚，硕士研究生
收稿日期：2008-10-22 **E-mail：**zj1979129@yahoo.com.cn

$\alpha(\cdot,\cdot)$ $(0 \le \alpha(\cdot,\cdot) \le 1)$。对任意一个组合 (x,x')，定义如下：

$$\begin{cases} p(x,x') = q(x,x')\alpha(x,x') & x \ne x' \\ p(x,x) = 1 - \int\limits_{x \ne x'} q(x,x')\alpha(x,x')\mathrm{d}x' & x = x \end{cases} \tag{1}$$

其中，$p(x,x')$为一个概率转移核，如果链在时刻 t 处于状态 x，即 $X^{(t)} = x$，则先由 $q(\cdot|x)$ 产生一个潜在的转移 $x \to x'$，然后根据概率 $\alpha(x,x')$接收 x' 作为马尔可夫链下一时刻的状态值，以概率 $1-\alpha(x,x')$拒绝转移到 x'，从而使链在下一时刻仍然处于状态 x。

本文目的是得到平稳分布的 $\pi(x)$，在给定 $q(\cdot,\cdot)$ 时，$\alpha(x,x') = \min\{1, \dfrac{\pi(x')q(x',x)}{\pi(x)q(x,x')}\}$，此时有

$$p(x,x') = \begin{cases} q(x,x') & \pi(x')q(x',x) \ge \pi(x)q(x,x') \\ q(x',x)\dfrac{\pi(x')}{\pi(x)} & \pi(x')q(x',x) < \pi(x)q(x,x') \end{cases} \tag{2}$$

3.2 最小二乘 Metropolis-Hasting(LSMH)算法

本节描述结合了 Metropolis-Hasting 抽样算法的定位算法。Metropolis-Hasting 抽样算法较独立，可以和多数定位算法相结合。最小二乘测量法是常见的定位方法，因此，本文以理想的最小二乘测量法为例，即假设测量距离是准确的，而参考节点的位置信息可能被攻击。

在现有基于测距的定位，如三边测量定位(Trilateration)或多边测量定位(Multilateration)中，通常采用基于最小二乘估计的算法。

理论上，如果一个非参考节点得到 2 个以上参考节点的位置信息及其到相应参考节点的距离，那么它就可以估计出自己的位置。得到的参考节点位置信息越多，估计出的位置越准确。若参考节点的信息被攻击，则计算出的位置是不准确的，该不准确程度取决于参考节点位置信息被攻击的程度和遭受攻击的参考节点个数。本文目的是提炼未被攻击的参考节点进行定位。由于较容易获得参考节点的原始分布特性，而被攻击后的位置统计特性能凭经验获得，因此利用 Metropolis-Hasting 取舍原则可以从被攻击的数据集里采集具有原始位置分布特性的数据。例如，假设在二维平面 $S_{L\times L}$ 上分布了 n 个无线传感器参考节点，它们的坐标分别为 $(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)$。在理想情况下，上述坐标各自服从随机均匀分布，即

$$\{(x,y) \mid x \sim U(-\frac{L}{2},\frac{L}{2}), y \sim U(-\frac{L}{2},\frac{L}{2})\} \tag{3}$$

因为横坐标与纵坐标之间、各个节点之间都是相互独立的，所以原始坐标的联合概率密度为

$$f(x,y) = (1/L)\times(1/L) = 1/L^2 \tag{4}$$

如果每个节点遭受攻击的概率为 P，则被攻击后的坐标服从正态分布

$$\{(x,y) \mid x \sim N(0,\delta^2), y \sim N(0,\delta^2), x,y \in [-\frac{L}{2},\frac{L}{2}]\} \tag{5}$$

被攻击后坐标的联合概率密度函数为

$$q(x,y) = (1-P)\cdot f(x)\cdot f(y) + P\cdot\exp\{-(x^2+y^2)/(2\delta^2)\}/(2\pi\delta^2) = (1-P)/L^2 + P\cdot\exp\{-(x^2+y^2)/(2\delta^2)\}/(2\pi\delta^2) \tag{6}$$

当 L 足够大时，$C \approx 1/(1-P)$。

采用 Metropolis-Hasting 抽样算法对非参考节点得到的参考节点位置信息进行处理，具体如下：

(1)获得 m 个参考节点的位置信息 $(x_1,y_1),(x_2,y_2),\cdots,(x_m,y_m)$，并产生 m 个随机变量 $r_1, r_2,\cdots,r_m \sim U(0,1)$。

(2)按概率 $q(\cdot|(x,y))$ 抽样得到(x',y')，即 $(x',y') \sim q(\cdot|(x,y))$。

(3)对于 i=1,2,…,m，如果有 $r_i < \min\{1, \dfrac{f(x',y')q((x',y'),(x,y))}{f(x,y)q((x,y),(x',y'))}\}$，则保留 (x',y')，否则舍弃 (x',y')。

当算法穷尽所有样本后，可以得出符合原始分布的样本，虽然无法保证被挑选出来的样本没有受到攻击，但实验结果表明，利用经过挑选的位置信息进行定位，其性能得到明显提升。

4 性能比较

通过 3 组实验比较原始最小二乘测量法 LS 和经过 Metropolis-Hasting 抽样改进后的最小二乘测量法 LSMH 的性能，包括平均定位误差和平均定位比例剩余。平均定位误差是所有节点的估测位置和真实位置之间相差的平均距离，在实验中，用节点的通信距离将其归一化。平均定位比例剩余是指在若干轮定位后，未得到定位的非参考节点占非参考节点总数的比例，它反映了算法的收敛速度。

本文在 Matlab 上实现了最小二乘测量法和改进后的最小二乘测量法 LSMH。它们仅在定位过程中运行一次，不考虑非参考节点在获得定位后再参与定位的情况。

本文不考虑少于 2 个参考节点参与定位的情况。在 LS 中，如果某个非参考节点听到的参考节点数目小于 2，那么该节点无法得到定位。而在 LSMH 中，若经过滤后的参考节点数小于 2，则无法得到定位。因此，本文认为缺少信息比信息错误更严重。

4.1 参数设置

本文实验参数设置如下：

(1)实验的默认设置如下：在 200 m×200 m 的正方形区域内随机均匀设置 200 个无线传感器节点，其中包括若干参考节点，这些参考节点的位置信息可能受到攻击。所有实验数据都是在独立运行 100 次之后得到的。

(2)所有节点的原始坐标按式(4)分布，其中，L=200。

(3)被感染的参考节点坐标服从式(6)分布，其中，正态分布的均值为 0，标准差为 δ。

(4)通过改变网络中的参考节点总数(Reference Number, RN)，可以调整非参考节点听到的平均参考节点数目。

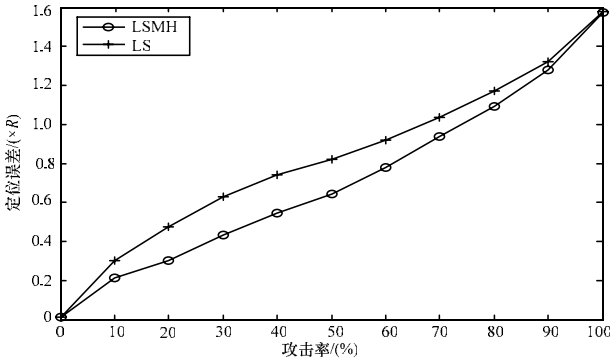
(5)参考节点被攻击的概率 TP(Tempered Percentage)在理想情况下为 0%，在最糟糕的情况为 100%。利用 MH 的目的是当攻击率介于 2 个极端情况之间时，可以尽可能筛选出有用的位置信息。

(6)节点的通信半径为 R，任何 2 个节点在通信半径内是连通的，在通信半径外是不连通的。

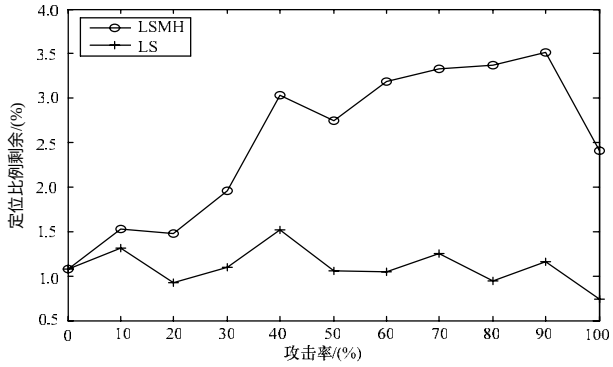
4.2 攻击率对定位算法性能的影响

讨论每个参考节点被攻击的概率对定位误差的影响。实验参数设置如下：参考节点个数为 50，通信半径为 50 m，标准差为 20 m，攻击率从 0%增加到 100%。

由图 1(a)可以看出，在攻击率增加的情况下，2 种算法的定位误差都迅速增加，但改进的 LSMH 算法始终维持着比 LS 低的误差，尤其是当攻击率处于 20%~60%之间时，LSMH 的误差比 LS 平均低 20%左右。由图 1(b)可以看出，LSMH 随着攻击率的增加造成参与定位的合格参考节点数目减少，导致越来越多的非参考节点无法完成定位，相对于 LS 来说，其定位比例剩余最高部分也处于攻击率在 20%~60%的阶段。



(a)定位误差



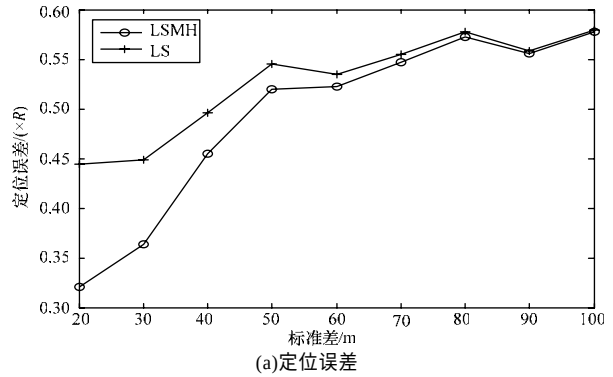
(b)定位比例剩余

图 1　攻击率对定位算法性能的影响

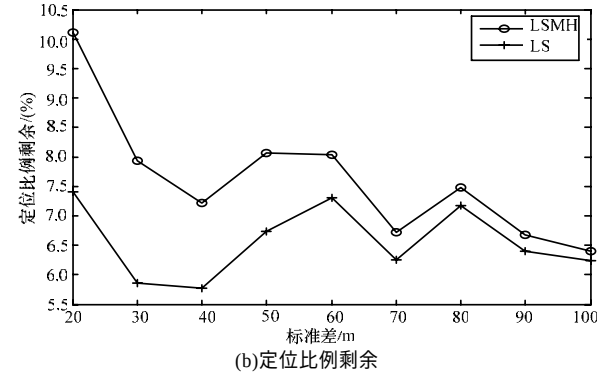
总体而言，在攻击率不断增加的情况下，LSMH 算法具有较好的过滤性能，保证了算法鲁棒性。

4.3　标准差对定位算法性能的影响

标准差的变化反应了被攻击后数据的变化范围。本文讨论在标准差从 20 m 增加到 100 m 的过程中，定位性能受到的影响。实验参数设置如下：参考节点个数为 30，通信半径为 50 m，攻击率为 20%。标准差对定位算法性能的影响见图 2。



(a)定位误差



(b)定位比例剩余

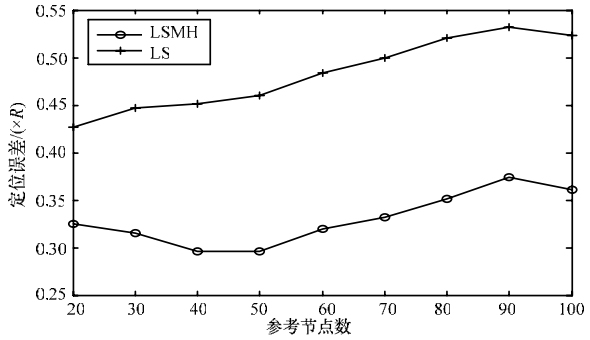
图 2　标准差对定位算法性能的影响

由图 2(a)可知，LSMH 和 LS 算法的定位误差都随着标准差的增加而变大，但 LSMH 的定位误差始终小于 LS，特别是当标准差较小时，LSMH 算法的优势更显著。当标准差增加到 70 m 以上时，2 种算法的误差曲线几乎相交。可见，当标准差大到一定程度时，LSMH 算法的优势将极大降低，辨识攻击数据的能力降低造成误差增加。在图 2(b)中，LS 算法的定位比例剩余略优于 LSMH 算法。随着标准差的增加，2 种算法的定位比例剩余越来越接近。在此期间，因为 LSMH 算法辨识能力降低，所以越来越多攻击信息被保留下来，造成定位比例剩余降低了近 4%。可见，LSMH 算法更适合在低标准差的情况下工作。

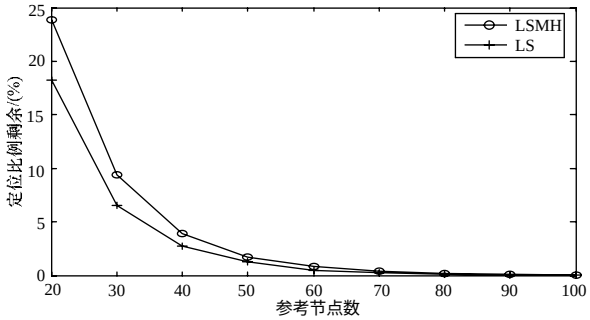
4.4　参考节点数对定位算法性能的影响

讨论当参考节点总数从 20 增加到 100 的过程中，定位算法性能受到的影响。实验参数设置如下：节点的通信半径为 50 m，攻击率为 20%，标准差为 20 m。

一般情况下，参考节点数目越多定位误差越小，但由图 3(a)可以看出，随着参考节点数目的增加，总体而言，LS 算法和 LSMH 算法的定位误差都有不同程度的增加，但后者较前者低 10%~17%。其原因是在攻击率一定的情况下，参考节点总数的增加带来了更多有用信息，但也带来了更多被攻击信息。LS 算法定位误差的增加表明增加的攻击信息发挥了比增加的有用信息更大的作用，而 LSMH 算法定位误差的增加表明其过滤性能在一定程度上和攻击率存在一个比例关系。图 3(b)表明，随着参考节点的增加，将有更多参考节点参与定位，从而使 2 种算法的定位比例剩余减少。



(a)定位误差



(b)定位比例剩余

图 3　参考节点数对定位算法性能的影响

5　结束语

本文介绍无线传感器网络在定位时可能遭受的威胁，利用 Metropolis-Hasting 抽样算法滤除被攻击的参考节点位置信息。下一步工作是将上述算法应用到已经存在的若干典型定位算法上，分析并比较它与不同算法结合后的性能差别。

(下转第 147 页)