LETTER Affinity Propagation Algorithm Based Multi-Source Localization Method for Binary Detection

Yan WANG^{†a)}, Long CHENG^{††}, Members, and Jian ZHANG^{†††}, Nonmember

SUMMARY Wireless sensor network (WSN) has attracted many researchers to investigate it in recent years. It can be widely used in the areas of surveillances, health care and agriculture. The location information is very important for WSN applications such as geographic routing, data fusion and tracking. So the localization technology is one of the key technologies for WSN. Since the computational complexity of the traditional source localization is high, the localization method can not be used in the sensor node. In this paper, we firstly introduce the Neyman-Pearson criterion based detection model. This model considers the effect of false alarm and missing alarm rate, so it is more realistic than the binary and probability model. An affinity propagation algorithm based localization method is proposed. Simulation results show that the proposed method provides high localization accuracy.

key words: wireless sensor network, binary detection, multi-source, localization, affinity propagation

1. Introduction

Wireless sensor network (WSN) with the advantages of low costs and energy savings has been rapidly promoted. It becomes a research hotspot technology in recent years. Each individual wireless sensor node has limited capacity of perception, receiving and sending information and computing power [1]. WSN not only consist of a large number of sensor node but also own the characteristics of concealment and fast and flexible deployment. Therefore, it can be used in complex terrain and difficult environments. The position information is critical for WSN applications such as geographic routing, data fusion and tracking. Therefore, the localization problem is one of the most important issues for WSN [2].

Since the computational complexity is high in the traditional source localization method, the source localization method for the binary sensor network is investigated in this paper. Two linear recursive three-dimensional source localization algorithms [3] are proposed by reorganizing the non-linear measurement equations into linear equations to improve the accuracy of localization. These methods take

Manuscript revised March 16, 2017.

Manuscript publicized May 10, 2017.

[†]The author is with the Department of Computer and Communication Engineering, Northeastern University, Qinhuangdao, 066–004 China.

^{††}The author is with the College of Information Science and Engineering, Northeastern University, Shenyang, 110–819 China.

^{†††}The author is with the School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu Province, P.R. China.

a) E-mail: ywang8510@gmail.com

DOI: 10.1587/transinf.2016EDL8235

into consideration the random sensor position noise. An optimization to maximum likelihood method [4] is proposed to estimate the position of acoustic source. The optimization to maximum likelihood algorithm that minimizes the sum of squares bias has been shown to consistently provide superior estimation performance. An iterative source localization method [5] is proposed. This method uses the anchor sensor to estimate the coarse source location. The posterior probability density function of the source is approximated using Monte Carlo method. A mean square error analysis method [6] is conducted to neglecting sensor position error and clock bias to improve the localization accuracy. A decentralized energy ratios based acoustic source localization method using the incremental gradient algorithm and normalized incremental gradient algorithm is proposed [7].

In this paper, an affinity propagation algorithm based multi-source localization method is proposed. This algorithm owns the lower computational complexity is proposed in the case of two sources. The optimal results can be obtained based on the estimation of affinity propagation algorithm. Simulation results show that the proposed method own higher localization accuracy and lower computational complexity.

2. Neyman-Pearson Detection Model

We consider a scenario where N sensor node and K acoustic sources are randomly deployed. The received signal strength for *i*th sensor node y_i can be expressed by

$$y_{i} = \begin{cases} \sum_{k=1}^{N} \frac{\beta_{k}}{d_{ik}^{\gamma/2}} + n_{i} , H_{1} \\ n_{i} , H_{0} \end{cases}$$
(1)

where H_1 denotes there are sources in the field and H_0 denotes no sources in the field, n_i is the measurement noise and it is modeled as zero mean Gaussian distribution with deviation σ_i , $\beta_k = \sqrt{S_k/2^{\gamma}}$, S_k is the energy of *k*th source, γ is the path loss coefficient, d_{ik} is the distance between *i*th sensor node and *k*th source.

The probability density function of measurement y_i under H_1 and H_0 can be obtained as:

$$P(y_i | H_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} \left(y_i - \sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2\right\} \quad (2)$$

<u>.</u>

Manuscript received December 7, 2016.

LETTER

$$P(y_i | H_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{y_i^2}{2\sigma^2}\right\}$$
(3)

According to Eqs. (2) and (3) and Neyman-Pearson criterion, the discriminatory equation is given by [8]

$$L_{i}(y_{i}) = \frac{P(y_{i} | H_{1})}{P(y_{i} | H_{0})}$$

= $\exp\left\{\frac{1}{2\sigma^{2}}\left(2y_{i}\sum_{k=1}^{N}\frac{\beta_{k}}{d_{ik}^{\gamma/2}} - \left(\sum_{k=1}^{N}\frac{\beta_{k}}{d_{ik}^{\gamma/2}}\right)^{2}\right)\right\} \stackrel{H_{1}}{\underset{H_{0}}{\overset{(4)}{\underset{H_{0}}{\underset{H_{0}}{\overset{(4)}{\underset{H_{0}}{\underset{H_{0}}{\overset{(4)}{\underset{H_{0}}{\underset{H_{0}}{\underset{H_{0}}{\overset{(4)}{\underset{H_{0}}{\underset{H_{0}}{\underset{H_{0}}{\underset{H_{0}}{\overset{(4)}{\underset{H_{0}$

where ω is the decision threshold value.

According to Eq. (4), we can obtain:

$$\frac{1}{2\sigma^2} \left(2y_i \sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}} - \left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}} \right)^2 \right)_{\substack{< \\ H_0}}^{H_1} \ln \omega$$
(5)

The Eq. (5) can be further converted as:

$$\underbrace{y_i \sum_{k=1}^{N} \frac{\beta_k}{d_{ik}^{\gamma/2}}}_{h} \stackrel{H_1}{\underset{H_0}{\overset{\sim}{\underbrace{\longrightarrow}}}} \sigma^2 \ln \omega + \frac{1}{2} \left(\sum_{k=1}^{N} \frac{\beta_k}{d_{ik}^{\gamma/2}} \right)^2 _{k}$$
(6)

According to the Eq. (6), we can obtain that

$$H_0: h \sim N\left(0, \sigma^2 \left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2\right)$$
(7)

$$H_1: h \sim N\left(\left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2, \sigma^2\left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2\right)$$
(8)

We set that
$$\mu_1 = \left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2$$
, $\sigma_1^2 = \sigma^2 \left(\sum_{k=1}^N \frac{\beta_k}{d_{ik}^{\gamma/2}}\right)^2$.
Therefore, the false alarm rate of the sensor node c

Therefore, the false alarm rate of the sensor node can be expressed:

$$P_F = P(h > k \mid H_0) = 1 - \Phi\left(\frac{k}{\sigma_1}\right)$$
(9)

where, $\Phi(\cdot)$ is the cumulative distribution function of standard normal distribution.

The detection rate of the sensor node is denoted as

$$P_D = P(h > k | H_1) = 1 - \Phi\left(\frac{k - \mu_1}{\sigma_1}\right)$$
(10)

According to Eq. (9), we can obtain $\frac{k}{\sigma_1} = \Phi^{-1}(1 - \alpha)$

and $\frac{\mu_1}{\sigma_1} = \frac{1}{\sigma} \sum_{k=1}^{N} \frac{\beta_k}{d_{ik}^{\gamma/2}}$, α is the false alarm rate.

The joint detection probability is shown as:

$$p_D = 1 - \Phi\left(\Phi^{-1}(1-\alpha) - \sum_{k=1}^N \frac{\beta_k}{\sigma} d_{ik}^{\gamma/2}\right)$$
(11)

3. Affinity Propagation Algorithm Based Localization Method

3.1 Basic Background of Affinity Propagation Algorithm

Affinity propagation (AP) algorithm is a clustering algorithm based on the message passing between the data [9], [10]. The main advantage of this algorithm is it does not require the number of clusters before running this algorithm in comparison with k-means method. The basic principle of the AP algorithm as follows: Let x_1 through x_n be a set of data points, and let *S* be a function that quantifies the similarity between any two points, such that $S(x_i, x_j) > S(x_i, x_k)$. The AP algorithm proceeds by alternating two message passing steps to update two matrices

- (1) The responsibility matrix \boldsymbol{R} which owns the values r(i, k) that quantify how well-suit x_k is to serve the exemplar for x_i , relative to other candidate exemplars for x_i .
- (2) The availability matrix A which has the values a(i, k). a(i, k) represents how appropriate it would be for x_i to pick x_k as its exemplar, considering other points' preference for x_k as an exemplar.

3.2 Proposed Method

The proposed method with the example of two sources is used to illustrate the process. It can be easily to extend to multi-source environment. Since the AP algorithm does not need to know the number of sources, it could estimate multiple cluster centers, i.e. the estimated number of sources may larger than the actual number of sources. The part of the estimated centers may deviate from the actual position of source. Therefore, we modify the classical AP algorithm to solve these problems. Firstly, the classical AP algorithm computes the initial positions of sources. Secondly, we propose a method to divide the estimated multiple position into two clusters. Finally, the centers of the new clusters are the final position of the sources. The steps of the proposed method as follows

Step 1. For a given data set which contains N elements, X_i and X_j are the different element in this set. We initialize the iteration number t = 1. We compute the similarity degree between X_i and X_j as follows:

$$s(i, j) = -||X_i - X_j||^2 \quad i \neq j$$
(12)

where, s(i, j) respects the similarity. If we can not obtain the prior knowledge, the self-similarity set as a constant. Normally, it is set as the mean of the input similar matrices as follows

$$s(l, l) = \frac{\sum_{i, j=1, i \neq j}^{N} s(i, j)}{N \times (N-1)}, \quad 1 \le l \le N$$
(13)

Step **2.** The Affinity Propagation algorithm proceeds by alternating two message passing steps, to update responsibility matrix and availability matrix. The responsibility matrix R has responsibility value r(i, j). The availability matrix A contains availability value a(i, j). Compute the responsibility value r(i, j) how well-suited x_j is to serve as the cluster center for x_i and the availability value a(i, j) as follows:

$$r(i, j) = s(i, j) - \max_{j' \neq j} \{a(i, j') + s(i, j')\}$$
(14)
$$a(i, j) = \begin{cases} \min\left\{0, r(j, j) + \sum_{i' \neq i, j} \max\{0, r(i', j)\}\right\} & i \neq j \\ \sum_{i' \neq j} \max\{0, r(i', j)\} & i = j \end{cases}$$
(15)

The availability value set as zero in the iterative initialization phase, i.e. a(i, j) = 0.

Step **3.** Compute the responsibility matrix *R* and availability matrix *A* as follows:

$$R_{t+1} = (1 - \lambda)R_t + \lambda R_{t-1} A_{t+1} = (1 - \lambda)A_t + \lambda A_{t-1}$$
(16)

where, $\lambda \in [0, 1]$ is the attenuation coefficient which is used to avoid the numerical oscillation in iteration.

If the ownership category of each sample remains the same in certain iteration or achieves the maximum iterations, the algorithm stops. Otherwise jump to the 2 step and update the number of iteration t, i.e. t = t + 1.

Step **4.** The maximum value of the sum of the responsibility matrix and availability matrix is the cluster center, it is expressed as:

$$c_i \leftarrow \underset{i \le j \le N}{\operatorname{argmax}} [r(i, j) + a(i, j)]$$
(17)

Step **5.** Compute the distance between the pair of two cluster centers

$$l = \sqrt{(x_c - x_c)^2 + (y_c - y_{c'})^2} \quad c \neq c'$$
(18)

where, (x_c, y_c) and $(x_{c'}, y_{c'})$ are the coordinates of the cluster center c_i and $c_{i'}$ respectively.

Step 6. Divide the initial positions in Eq. (17) according to Eq. (18). In order to illustrate this step, we provide a simple example using two sources. As shown in Fig. 1, there are six initial estimated positions according to step 4. The distance between center c_1 and c_4 is the furthest distance. Therefore, we divide the initial positions into two clusters according to center c_1 and c_4 . Since the center c_2 is nearer c_1 in comparison with center c_1 , so the c_2 belongs to c_1 .

4. Simulation Results

In this section, we evaluate the performance of the proposed method. N sensor nodes and K sources are randomly deployed in the monitoring field. The default parameters in the



Fig. 1 An example of Step 6.

Table 1The default parameters.

Parameters	Symbols	Default Values
The number of sensor nodes	N	200
The number of sources	Κ	2
The energy of <i>k</i> th source	S_k	3000
The standard deviation of measurement noise	σ_{i}	1
The false alarm rate	α	0.02



Fig. 2 The localization error versus the standard deviation of measurement noise.

simulation experiments are shown in Table 1. We employ the average localization error to evaluate the performance of the proposed method:

Error =
$$\frac{1}{N_t K} \sum_{i=1}^{N_t} \sum_{k=1}^{K} \sqrt{(\hat{x}_{ki} - x_k)^2 + (\hat{y}_{ki} - y_k)^2}$$
 (19)

where, $N_t = 200$, (x_k, y_k) is the true position for *k*th source, $(\hat{x}_{ik}, \hat{y}_{ik})$ is the estimated position of *k*th source for *i*th experiment.

Figure 2 shows the relationship between the localization error and the standard deviation of measurement noise under different number of sensor nodes. It can be observed that the localization error increases with the standard deviation of measurement noise increase. This is because the larger noise induces stronger interference. The localization error decreases with the number of sensor nodes increases.

Figure 3 shows the impact of the packet loss rate (PLR) on the localization error under different energies of source. It indicates that the localization error of the proposed method remains unchanged with the increase of PLR. So, the proposed method is robust to packet loss rate. It can



Fig. 3 The localization error versus the packet loss rate.

Table 2 Running times.

The number of sensor nodes	Proposed Method	Traditional Method
N=200	0.21168 s	8.168402 s
<i>N</i> =300	0.38840 s	15.797392 s
<i>N</i> =400	1.59237 s	24.000729 s



Fig.4 The localization error versus the false alarm rate.

mitigate the effect of PLR. With the increase of the energy of source, the localization error increases. This is because the larger energy source will induce larger interfering with each source.

Table 2 shows the running times of the proposed method and the traditional method. The traditional method uses the particle swarm optimization method to estimate the position of sources. The two methods are coded using Matlab 2008 and tested on a Windows 7 Professional workstation with Intel(R) Core(TM) 2 Duo CPU E7300 @ 2.66GHz and 2.00GB RAM. Compared with the traditional method, the proposed method obviously owns lower running time. Therefore, the proposed method has lower computational complexity.

Figure 4 shows the relationship between the false alarm rate and the localization error. It can be observed that the localization error increase with the increase of false alarm rate. This is because the larger false alarm rate will result in the number of false alarm nodes increases. And the average localization error increases as the number of sources *K* increases. This is because the interaction effects make localization performance worse if there are too many sources.

5. Conclusion

In this paper, we investigate the multi-source localization method based on binary detection for wireless sensor network. The detection model is firstly introduced. The Neyman-Pearson detection which considers the false alarm rate and missing rate is presented. The affinity propagation algorithm based localization method is proposed to estimate the initial positions of sources. The further processing steps are proposed to improve the localization accuracy. Simulation results show that the proposed method achieves high localization accuracy.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant No. 61403068, No. 61232016 and No. U1405254; Natural Science Foundation of Hebei Province under Grant No. F2015501097 and No. F2016501080; the PAPD fund and Fundamental Research Funds for the Central Universities of China (N152302001).

References

- P. Guo, J. Wang, B. Li, and S. Lee, "A variable threshold-value authentication architecture for wireless mesh networks," Journal of Internet Technology, vol.15, no.6, pp.929–936, June 2014.
- [2] I. Amundson, J. Sallai, X. Koutsoukos, and A. Ledeczi, "Mobile sensor waypoint navigation via RF-based angle of arrival localization," International Journal of Distributed Sensor Networks, Article ID 842107, 2012.
- [3] X. Qu and L. Xie, "Recursive source localisation by time difference of arrival sensor networks with sensor position uncertainty," IET Control Theory and Applications, vol.8, no.18, pp.2305–2315, 2014.
- [4] M.Z. Rahman, D. Habibi, and I. Ahmad, "Source localisation in wireless sensor networks based on optimised maximum likelihood," Australasian Telecommunication Networks and Applications Conference, pp.235–239, 2008.
- [5] E. Masazade, R. Niu, P.K. Varshney, and M. Keskinoz, "Energy aware iterative source localization for wireless sensor networks," IEEE Trans. Signal Process., vol.58, no.9, pp.4824–4835, 2010.
- [6] Y. Wang, J. Huang, L. Yang, and Y. Xue, "TOA-based joint synchronization and source localization with random errors in sensor positions and sensor clock biases," Ad Hoc Networks, vol.27, pp.99–111, 2015.
- [7] S. Wang and J. Wang, "Decentralized acoustic source localization with unknown source energy in a wireless sensor network," Measurement Science & Technology, vol.18, no.12, pp.3768–3776, 2007.
- [8] Y. Yang, R.S. Blum, and B. Sadler, "Energy-efficient routing for signal detection in wireless sensor networks," IEEE Trans. Signal Process., vol.57, no.6, pp.2050–2063, 2009.
- [9] B.J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol.315, no.5814, pp.972–976, 2007.
- [10] D.-W. Chen, J.-Q. Sheng, J.-J. Chen, and C.-D. Wang, "Stabilitybased preference selection in affinity propagation," Neural Computing & Applications, vol.25, no.7, pp.1809–1822, 2014.