ALGORITHM OF SIGNAL SOURCE LOCALIZATION VIA SENSOR NETWORKS

Lianming Sun and Xinyu Liu

Faculty of Environmental Engineering
The University of Kitakyushu
1-1 Hibikino, Wakamatsu-ku, Kitakyushu 808-0135, Japan
sun@kitakyu-u.ac.jp

Abstract. *The location information of a signal source is localized through a sensor network in many practical applications. Besides the measurement noise, sometimes the measured data are contaminated with multi-path interferences, which may distort the main-lobe of the correlation functions and lead to localization errors. In this paper the localization algorithm of the signal source under the multi-path environment is considered. A refinement for the time difference estimation of the signal arrival times between the different sensors is applied to enhancing the main-lobe of the correlation functions, and the mutual error verification is further introduced to reduce the affection caused by the estimation errors of time difference. The effectiveness is illustrated through the simulation examples of the low frequency seismograms.*
Keywords: Source localization, Estimation of time difference, Error verification, Low frequency seismograms

1. **Introduction.** Localization of a signal source is a fundamental problem in many application fields. For example, the location information of the targets is used in radar or ultrasonic wave recognition to determine the next action in the instrumentation unit of autonomous vehicles, autonomous robots, the location information of the epicenter is necessary in seismic analysis and early warning system [1, 2, 3, 4], the fault detection is used for monitoring the underground pipes or cables [5], and the beam forming is used in the transmitters/receivers in communication and acoustic systems [6]. Consequently, a lot of localization systems and signal/data processing techniques have been proposed. Most of the conventional methods deal with the localization problem in two steps. In the first step, they estimate the time difference of the signal wave arriving at the different sensors, while the others estimate the arrival directions to make use of the space diversity under the case where the signal waves have propagated through a long distance. In the second step, the source position is deduced using a localization algorithm. In the existing methods, the time difference can be obtained by maximizing the correlation functions of the measured signals, or maximizing the semblance function [7, 8], or synchronizing the peak of correlation functions of the detected signal waves such as the seismic waves in the array record of seismograms in time domain [9, 10]. Moreover, it is shown that optimizing the semblance function in the frequency domain can improve the computational efficiency of the processing algorithm [11]. On the other hand, the arrival directions are estimated through some subspace methods based on array signal processing techniques in radar and communication systems [12, 13], and the performance is analyzed [14]. Compared with the sub-optimal two-step methods since the fact of each arrival signal coming from a same source is not considered in the first step, the direct localization methods using the centralized one-step algorithms are also investigated in an active radar network [15].

In many of the existing methods, it is assumed that the signal waves propagate to the sensors or receivers directly through a single path, or the multi-path interferences do not affect the estimation of location too much when using some instinctive characteristics such as space diversity. Nevertheless, in some severe propagation environment, the measured signals are contaminated with the multi-path interferences caused by wave reflection or scattering, and the performance of the source location's estimation degrades largely due to the affection of the side-lobes associated with the multi-path interferences. Furthermore, the localization performance degrades largely if the source signal has severe band limitation where there are insufficient frequency components for time difference estimation or direction of arrivals.

On the other hand, if the data obtained by each sensor in the sensor network can be effectively fused, much information may be detected from the measurements and it may help to reduce the affection of measurement noise as well as the multi-path interferences. In this paper, a new two-step localization algorithm using the data measured by the sensor network is presented. It introduces a time refinement to the estimation of the time difference between the signals measured by different sensors, and fuses the information detected from each sensor through a weight matrix, which is determined by the result of mutual error verification. Then, the localization of the signal source turns to a nonlinear optimization problem under the multi-path environment where the source signal is band limited. The effectiveness of the proposed algorithm is also investigated through a numerical simulation example of the epicenter location from the low frequency seismograms.

The rest of the paper is organized as follows. In the next section, the main description of the problem considered in this paper is summarized. In Section 3, the estimation algorithm for the time difference between the measured signals is investigated, then the precision evaluation of time difference estimation is illustrated, and the information fusion through a weight matrix and the nonlinear optimization to localize the source are shown in Section 4. Section 5 demonstrates a numerical simulation example. Finally, the conclusion and the future research work are given in Section 6.

2. **Problem Statement.** The diagram of sensor network considered in this paper is illustrated in Figure 1. In the network with L sensor elements, the signal at the location of a signal source is denoted as $s(k)$, which cannot be measured directly. Here k indicates a normalized instant of kt_s, and t_s is the sampling interval of the measured signals.
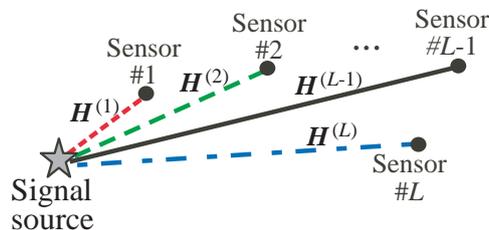

FIGURE 1. Illustration of sensor network and signal source

$s(k)$ propagates from the source location to the lth sensor through the channel $H^{(l)}$, where $l = 1, \ldots, L$. Let the signal measured by the lth sensor be denoted as $a^{(l)}(k)$, which is contaminated with the noise term $n^{(l)}(k)$. Then, the relation of the measured signal $a^{(l)}(k)$ at the lth sensor and $s(k)$ can be approximated by

$$a^{(l)}(k) = h_0^{(l)} s\left(k - \tau_0^{(l)}\right) + h_1^{(l)} s\left(k - \tau_1^{(l)}\right) + \cdots + h_M^{(l)} s\left(k - \tau_M^{(l)}\right) + n^{(l)}(k), \qquad (1)$$

where $\tau_0^{(l)}$, $\tau_1^{(l)}$, $\ldots$ denote the delay time related with the distance of signal wave propagation, while $h_0^{(l)}$, $h_1^{(l)}$, $\ldots$ are the attenuation coefficients. In the multi-path environment,

besides the main wave $h_0^{(l)}s\left(k - \tau_0^{(l)}\right)$, there are several sub-waves $h_1^{(l)}s\left(k - \tau_1^{(l)}\right),\ldots,$ $h_M^{(l)}s\left(k - \tau_M^{(l)}\right)$, which will alter the waveforms, and affect the estimation of source location. Moreover, in some practical applications, the source signal $s(k)$ is such a band limited signal that the estimation of the channel parameters, or the channel identification is not an easy task.

Many conventional two-step localization methods attempt to estimate the time difference of the main wave $\tau_0^{(l_1,l_2)}$ between the signal measured by the l_1th and the l_2th sensors. The estimation can be performed through the correlation functions, the semblance of the signal waveforms, the difference of waveforms, etc. However, under the multi-path environment, the sub-waves will generate side-lobe in the correlation function. The distortion of the main-lobe caused by the side-lobe leads to some considerable estimation errors that may yield large mismatch of the source location. Consequently, the affection of the side-lobe should be reduced, and the estimation errors of time difference are also dealt with to improve the localization performance in the new algorithm.

3. **Estimation of Time Difference.** A new two-step algorithm is proposed to localize the signal source. To decrease the affection of noise terms and the affection of the side-lobe caused by the multi-path interferences, the correlation functions are calculated from the average periodograms of the measured signals first, and then the estimation is updated through refining the periodograms in the first step of the proposed algorithm.

3.1. **Initial estimation using periodograms.** Let the discrete-time Fourier transform (DFT) of the measured signal in the ith time window $[t_i, t_i + N - 1]$ be denoted as

$$A_i^{(l)}\left(e^{jn\omega}\right) = \sum_{k=0}^{N-1} a^{(l)}(t_i + k)e^{-jnk\omega}, \tag{2}$$

where t_i is the time offset of the ith window, N is the window length, while ω is the frequency interval $2\pi/N$. If N is chosen as a power of 2, the discrete-time Fourier transform can be performed through fast Fourier transform (FFT) algorithm to reduce the computation load. Furthermore, let the average periodogram of $a^{(l_1)}(k)$ and $a^{(l_2)}(k)$ be given by shifting the time window to $t_0,\ldots,t_{I-1}$,

$$\bar{A}^{(l_1,l_2)}\left(e^{jn\omega}\right) = \frac{1}{I}\sum_{i=0}^{I-1}\left(A_i^{(l_1)}\left(e^{jn\omega}\right)\right)^* A_i^{(l_2)}\left(e^{jn\omega}\right), \tag{3}$$

where $*$ indicates the complex conjugate.

Correspondingly, the correlation function defined in the time domain

$$R^{(l_1,l_2)}(\tau) = \frac{1}{K}\sum_{k=1}^{K} a^{(l_1)}(k_0 + k)a^{(l_2)}(k_0 + k + \tau) \tag{4}$$

can be approximated by

$$R^{(l_1,l_2)}(\tau) \approx \frac{1}{N}\sum_{n=0}^{N-1} \bar{A}^{(l_1,l_2)}\left(e^{jn\omega}\right)e^{jn\tau\omega} \tag{5}$$

in the frequency domain. Then the initial time difference between $a^{(l_1)}(k)$ and $a^{(l_2)}(k)$ can be given by

$$\hat{\tau}_{\text{init},0}^{(l_1,l_2)} = \arg\max_{\tau} R^{(l_1,l_2)}(\tau) \tag{6}$$

from the periodogram as in (5). If the distortion caused by multi-path interferences can be neglected, the estimation in (6) might be a good estimation for $\tau_0^{(l_1,l_2)} = \tau_0^{(l_2)} - \tau_0^{(l_1)}$. Nevertheless, the distortion may shift position of main-lobe of the correlation functions so

the estimation in (6) may have large errors, which should be reduced further to guarantee the effectiveness of source localization.

3.2. **Refinement of time difference estimation.** The affection of the measurement noise and the side-lobe caused by multi-path interferences is considered in this section. In the conventional methods, the cepstral prefiltering that utilizes inverse of the spectra is often used to reduce the side-lobe. However, if the source signal has severe band limitation, the performance of cepstral prefiltering degrades at the frequency points where there are not sufficient frequency components. In order to improve the estimation performance, a positive number c is introduced into the refinement of periodogram as follows:

$$\bar{A}_{\text{ref}}^{(l_1,l_2)}\left(e^{jn\omega}\right) = \frac{\bar{A}^{(l_1,l_2)}\left(e^{jn\omega}\right)e^{-jn\hat{\tau}_{\text{init},0}^{(l_1,l_2)}\omega}}{\left(\bar{A}^{(l_1,l_1)}\left(e^{jn\omega}\right)+c\right)^{\alpha}\left(\bar{A}^{(l_2,l_2)}\left(e^{jn\omega}\right)+c\right)^{\alpha}}, \tag{7}$$

where α and c are the parameters determined by the spectra of noise terms and the source components. To reduce the influence of noise terms, the parameter c is such a positive real number that the magnitude of $\bar{A}_{\text{ref}}^{(l_1,l_2)}\left(e^{jn\omega}\right)$ at the frequency points outside the band limitation remains small, and hence it can be selected with respect to the magnitude range and the magnitude boundary between the noise and the source signal components. Moreover, α is a non-negative constant $\alpha = 1/2 - \beta \le 1/2$, and $0 \le \beta \le 1/2$. (7) can also be considered as a weight function $\left(\bar{A}^{(l_1,l_1)}\left(e^{jn\omega}\right)+c\right)^{\beta}\left(\bar{A}^{(l_2,l_2)}\left(e^{jn\omega}\right)+c\right)^{\beta}$ that is introduced into the numerator of the common cepstral prefiltering to reduce the influence of noise. When $\beta = 0$, $\alpha = 1/2$, (7) becomes similarly as the common cepstral prefiltering where the denominator can cancel a large part of the side-lobe of the numerator under low noise environment. For $0 < \alpha, \beta < 1/2$, $\bar{A}_{\text{ref}}^{(l_1,l_2)}\left(e^{jn\omega}\right)$ might have lower side-lobe than that of $\bar{A}^{(l_1,l_2)}\left(e^{jn\omega}\right)$. On the contrary, if the measured data contain large noise, $\beta = 1/2$, $\alpha = 0$ will yield the spectrum without any refinement. Therefore, the choice of α and β can be a compromise between decreasing the effect of side-lobe and the influence of noise.

The refinement of the time difference estimation is calculated as

$$\hat{\tau}_{\text{ref},0}^{(l_1,l_2)} = \arg\max_{\tau} R_{\text{ref}}^{(l_1,l_2)}(\tau), \tag{8}$$

where the refinement of the correlation functions

$$R_{\text{ref}}^{(l_1,l_2)}(\tau) \approx \frac{1}{N}\sum_{n=0}^{N-1}\bar{A}_{\text{ref}}^{(l_1,l_2)}\left(e^{jn\omega}\right)e^{jn\tau\omega}. \tag{9}$$

Consequently, the estimation of time difference is updated by

$$\hat{\tau}_0^{(l_1,l_2)} = \hat{\tau}_{\text{init},0}^{(l_1,l_2)} + \hat{\tau}_{\text{ref},0}^{(l_1,l_2)}. \tag{10}$$

4. **Localization of Signal Source.** In this section, the second step of the algorithm to localize the signal source is investigated by using the time difference estimation given in Section 3. Although the estimation errors caused by the measurement noise and multi-path interferences have been decreased by the refinement of periodograms, a few errors still remain in the estimation due to the insufficient frequency components in the band limited signals. To guarantee effective localization, the errors are mutually verified before using the estimation of time difference to localize the signal source.

4.1. **Mutual error verification.** After estimating the time difference $\hat{\tau}_0^{(l_1,l_2)}$ for $l_1, l_2 = 1, 2, \ldots, L$, some redundancy information can be used to evaluate the estimation. Notice that for the true time difference, the following property

$$\tau_0^{(l_1,l_2)} = \tau_0^{(l_2)} - \tau_0^{(l_1)} = \tau_0^{(l,l_2)} - \tau_0^{(l,l_1)} \tag{11}$$

holds for $l, l_1, l_2 = 1, 2, \ldots, L$. Then, the estimation of $\hat{\tau}_0^{(l_1, l_2)}$ can be evaluated by verifying the difference between $\hat{\tau}_0^{(l_1, l_2)}$ and $\hat{\tau}_0^{(l, l_2)} - \hat{\tau}_0^{(l, l_1)}$, where $\hat{\tau}_0^{(l, l_2)}$ and $\hat{\tau}_0^{(l, l_1)}$ are estimated respectively by (10). Define the error index for the time difference estimation by

$$\varepsilon^{(l_1, l_2)} = \sum_{l \neq l_1, l_2} \left| \hat{\tau}_0^{(l_1, l_2)} - \left(\tau_0^{(l, l_2)} - \tau_0^{(l, l_1)} \right) \right|. \tag{12}$$

It is seen that the larger $\varepsilon^{(l_1, l_2)}$ is, the more errors exist in $\hat{\tau}_0^{(l_1, l_2)}$. Therefore, the matrix, whose (l_1, l_2)th element is $\varepsilon^{(l_1, l_2)}$, can be considered as indices of all the estimated time difference, and it can be used to determine the weight coefficients in the estimation of the source location.

4.2. **Estimation of source location.** Using the estimation of $\hat{\tau}_0^{(l_1, l_2)}$, and the corresponding error index $\varepsilon^{(l_1, l_2)}$ for $l_1, l_2 = 1, 2, \ldots, L$, the source location can be estimated by solving the following optimization problem

$$\arg \min_{x, y, z} \sum_{l_1, l_2 = 1}^{L} w^{(l_1, l_2)} \left(\Delta^{(l_1, l_2)} - v \hat{\tau}_0^{(l_1, l_2)} \right)^2, \tag{13}$$

where $w^{(l_1, l_2)}$ is the weight coefficient with respect to the estimation error $\varepsilon^{(l_1, l_2)}$, x, y, z are the coordinates of the source location, $\Delta^{(l_1, l_2)}$ is the distance difference between the distance from the l_2th sensor, and the distance from the l_1th sensor to the source location as follows:

$$\Delta^{(l_1, l_2)} = \left| \left[\begin{array}{c} x \\ y \\ z \end{array} \right] - \left[\begin{array}{c} x^{(l_2)} \\ y^{(l_2)} \\ z^{(l_2)} \end{array} \right] \right| - \left| \left[\begin{array}{c} x \\ y \\ z \end{array} \right] - \left[\begin{array}{c} x^{(l_1)} \\ y^{(l_1)} \\ z^{(l_1)} \end{array} \right] \right|, \tag{14}$$

where $|\cdot|$ indicates the distance, $x^{(l)}$, $y^{(l)}$, $z^{(l)}$ are the coordinate of the lth sensor, and v is the propagation velocity of the signal wave. Here the weight coefficient is given by

$$w^{(l_1, l_2)} = e^{\frac{1 - \varepsilon^{(l_1, l_2)}}{d}}, \tag{15}$$

where d is a positive number, which is determined by the range of $\varepsilon^{(l_1, l_2)}$ and the sensor number. Generally, d is set as a small value for the small range of $\varepsilon^{(l_1, l_2)}$ when the sensor number L is large, and vice versa.

5. **Numerical Simulation Example.** In the numerical example, the low frequency seismograms related to low frequency seismic ground tremors are considered. The tremors occur with the moments released at deep subduction plate interfaces, and have some characteristic motions related to the subduction zone and active fault. Detection, localization of the epicenter and analysis on these low frequency seismograms help us to explore the seismic motions of deep ground structure, or to evaluate the performance of earthquake-proof for deep ground structure. Therefore, localization the epicenter is an essential problem in the low frequency seismograms. Different from the ordinary seismograms, the low frequency seismograms have not explicit P-wave and S-wave arrivals, so the time difference between the signal measured at several tiltmeters in an earthquake monitoring network [1].

Let the location of the first tiltmeter station be the reference position, and then true locations of epicenter in each simulation run are randomly distributed as normal distribution of $\mathcal{N}(39.2899', (3.2909')^2)$ west in longitude, $\mathcal{N}(20.1193', (2.6903')^2)$ south in latitude, and uniform distribution as $10 \sim 15$km in the depth from the reference position. Moreover, there is a refraction wave from the geological layer with depth of $\mathcal{N}(21, 1)$km. The main frequency components in $s(k)$ are within $0.03 \sim 1.1$Hz, so the measured signals $a^{(l)}$ are also low frequency band limited signals.

There are 10 tiltmeter stations in the sensor network. The distance between every two stations is normally distributed among 18.1km $\sim$ 20.5km. The seismogram records are sampled at the sampling rate of 20Hz, and the records within a time window of 225s are used for the data analysis. In the example, the signal to noise ratio is set as 15dB.

An example of the normalized magnitude spectrum of the measured signal is illustrated in Figure 2, where the magnitude range is 8 $\sim$ 700. It is seen that the frequency components are concentrated within a narrow low frequency band, and the components whose magnitude is lower than 15 are the noise, so the parameters of the propagation channel are not estimated easily using the band limited signals if little of the channel information can be available.

In the algorithm, the data window length is chosen as $N = 1024$ to perform FFT. The constant number c is chosen as 50 with respect to suppression of the magnitude of the noise components, β is given by $1/8$ so $\alpha = 3/8$ in (7), whereas d in (15) is 3 with respect to the range of $0 \sim 12$ for $\varepsilon^{(l_1, l_2)}$ in this numerical example. The simulation of localization is performed for 30 simulation runs, where the source location has slight variation in every simulation run. The average of estimated coordinates of the source location as well as the standard deviation are summarized in Table 1, where the estimates are close to the true ones.

As a comparison, the results obtained by the conventional methods without refinement and mutual error verification have large estimation errors, which are mainly caused by the multi-path interferences and signal band limitation.
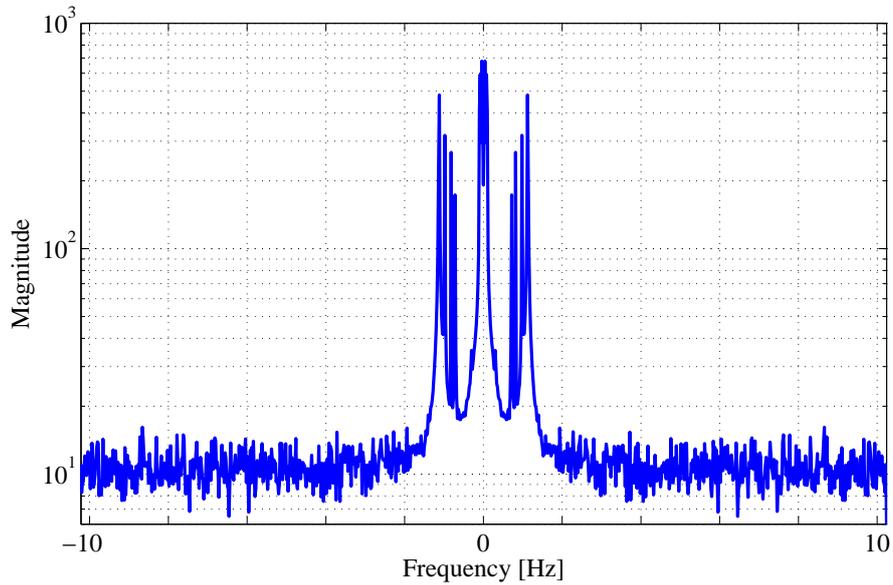

FIGURE 2. Example of magnitude spectrum of the measured signal

TABLE 1. Estimation of epicenter locations

Location to reference position	True values	Proposed method	Conventional method
Depth[km]	12.8018	12.5000	9.0776
Standard deviation		± 2.2146	± 6.3705
West in longitude [']	39.2899	39.1985	36.5919
Standard deviation		± 0.6261	± 5.1950
South in latitude [']	20.1193	20.0855	18.7129
Standard deviation		± 0.3608	± 2.6566

6. **Conclusions.** The problem of signal source localization has been investigated through a sensor network. It has been shown that by introducing the refinement for time difference estimation, and by introducing weight coefficients determined by the mutual error verification into the nonlinear optimization problem, the proposed algorithm can work under the multi-path environment. The effectiveness of the algorithm has been demonstrated by the analysis of low frequency seismograms. The algorithm for the more complex environment such as multi-source or severe multi-path interferences and the application to the real seismograms will be investigated in the future work.

REFERENCES

[1] K. Obara, Nonvolcanic deep tremor associated with subduction in southwest Japan, *Science*, vol.269, pp.1679-1681, 2002.

[2] H. Hirose and K. Obara, Repeating short- and long-term slow slip events with deep tremor activity around the Bungo channel regions, southwest Japan, *Earth Planets Space*, vol.57, pp.961-972, 2005.

[3] Committee for Earthquake Prediction Studies, Seimological Society of Japan, *The Science of Earthquake Prediction*, University of Tokyo Process, 2007.

[4] S. Saha, D. Mukherjee and S. Mukhopadhyay, Online detection and location estimation of earthquake events using continuous wavelet transform, *IEEE International Conference on Control, Measurement and Instrumentation*, pp.77-82, 2016.

[5] S. Lee et al., Fault detection of pipeline in water distribution network system, *International Journal of Mathematical, Computational, Physical, Electrical and Computer Engineering*, vol.6, no.4, pp.490-495, 2012.

[6] A. Alaa and L. Ray, Design and performance of a real-time acoustic beam forming system, *Proc. of IEEE Sensors*, pp.1-4, 2013.

[7] N. S. Neidell and M. T. Taner, Semblance and other coherency measures for multichannel, *Geophysics*, vol.36, pp.482-497, 1971.

[8] Y. Asano, K. Obara and Y. Ito, Spatiotemporal distribution of very-low frequency earthquakes in Tokachi-oki near the junction of the Kuril and Japan trenches revealed by using array signal processing, *Earth Planets Space*, vol.60, pp.871-875, 2008.

[9] C. Shirakawa and T. Iwata, Ground motion characteristics of South-East Kyoto Basin site using three dimensional small aperture seismic array at Uji Campus, Kyoto University, *Annuals of Disaster Prevention Research Institute, Kyoto University*, no.50(B), pp.251-258, 2007.

[10] H. Miura and S. Midorikawa, Effects of 3-D deep underground structure on characteristics of rather long-period ground motion, *Earthquake*, no.54, pp.381-395, 2001.

[11] L. Sun, W. Zhu and X. Zhu, Epicenter location via genetic algorithm for low frequency seismograms, *ICIC Express Letters*, vol.6, no.4, pp.953-958, 2012.

[12] B. Rao and K. Hari, Performance analysis of root-MUSIC, *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.12, pp.1939-1949, 1989.

[13] R. Roy and T. Kailath, ESPRIT – Estimation of signal parameters via rotational invariance techniques, *IEEE Trans. Acoust., Speech, Signal Processing*, vol.37, no.7, pp.984-995, 1989.

[14] J. Delmas, M. Korso, H. Gazzah and M. Castella, CRB analysis of planar antenna array for optimizing near-field source localization, *Signal Processing*, vol.127, no.1, pp.117-134, 2016.

[15] J. Bosse, O. Krasnov and A. Yarovoy, Direct target localization with an active radar network, *Signal Processing*, vol.127, no.1, pp.21-35, 2016.