ANLoC: An Anomaly-Aware Node Localization Algorithm for WSNs in Complex Environments

Pengfei Xu [1], Tianhao Cui [1] and Lei Chen [1,2,3,*]

[1] School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 1217043202@njupt.edu.cn (P.X.); 1018041227@njupt.edu.cn (T.C.)

[2] Jiangsu Key Laboratory of Big Data Security & Intelligent Processing, Nanjing University of Posts &Telecommunications, Nanjing 210023, China

[3] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

* Correspondence: chenlei@njupt.edu.cn

check for updates

Abstract: Accurate and sufficient node location information is crucial for Wireless Sensor Networks (WSNs) applications. However, the existing range-based localization methods often suffer from incomplete and detorted range measurements. To address this issue, some methods based on low-rank matrix recovery have been proposed, which usually assume noises follow single Gaussian distribution or/and single Laplacian distribution, and thus cannot handle the case with wider noise distributions beyond Gaussian and Laplacian ones. In this paper, a novel Anomaly-aware Node Localization (ANLoC) method is proposed to simultaneously impute missing range measurements and detect node anomaly in complex environments. Specifically, by utilizing inherent low-rank property of Euclidean Distance Matrix (EDM), we formulate range measurements imputation problem as a Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) model, where complex noise is fitted by Mixture of Gaussian (MoG) distribution, and node anomaly is sifted by $\ell_{2,1}$-norm regularization. Meanwhile, an efficient optimization algorithm is designed to solve proposed RRMD model based on Expectation Maximization (EM) method. Furthermore, with the imputed EDM, all unknown nodes can be easily positioned by using Multi-Dimensional Scaling (MDS) method. Finally, some experiments are designed to evaluate performance of the proposed method, and experimental results demonstrate that our method outperforms three state-of-the-art node localization methods.

Keywords: wireless sensor networks; anomaly-aware node localization; low-rank matrix decomposition; mixture of Gaussians

1. Introduction

Wireless Sensor Networks (WSNs) have made great progress and have been widely used in various fields, such as environmental monitoring, intelligent transportation, target tracking, etc. [1–3]. The premise that these applications work well is accurate location information acquisition [4,5]. Up to now, many localization methods have been proposed, which could be divided into two categories [6]. One is called range-based localization method that could achieve more accurate positioning, but the computation and communication overhead is large while some hardware support is also required. The other is named range-free localization method that is generally suitable for low power and cost applications, yet their positioning accuracy is low.

In this paper, we mainly focus on the range-based localization methods, which could be described as follows: in WSNs applications, some sensor nodes are randomly disposed, and a few of them called anchors could get actual location information by GPS device or other equipment. Then, with

the pair-wise range measurements between nodes and actual location of anchors, the location of all unknown nodes could be easily estimated. In general, range-based localization methods often depend on a large amount of accurate inter-node distance information. However, in practical situations, limited by the energy of sensors or the distribution of nodes in application scenario, only a small amount of inter-node distance information could be obtained [7]. Additionally, due to complex environments and uncertain hardware abnormality, the range measurements inevitably suffer from some errors, which lead to lower positioning accuracy. Usually, these errors can be regarded as a mixture of complex noise and anomaly. In reference [6], the complex noise is considered to be caused by environmental interference, malicious attacks, or other unpredictable factors, and is assumed to be a mixture of Gaussian noise with single distribution and outlier noise with single distribution. Correspondingly, Xiao et al. [6] proposed a method based on low-rank matrix completion to sift the complex noise by adopting both Frobenius-norm regularization and ℓ_1-norm regularization. However, presetting the complex noise into two known noise types was too arbitrary. Actually, the noise distribution in practical applications was often unknown and showed a wider noise distribution beyond Gaussian and Laplacian ones. On the other hand, reference [8] handled one more general application scenario with the co-existence of complex noise and anomaly nodes, where anomaly nodes were defined as ones with abnormal transmission module or unpredictable hardware defects. In reference [8], a $\ell_{2,1}$-norm regularization term was employed to detect the anomaly from the corrupted range measurements. However, similar to reference [6], the so-called complex noise is still limited to be modeled as two known noise types, i.e., the mixture of outlier noise and Gaussian noise.

To address this limitation, we propose a novel Anomaly-aware Node Localization (ANLoC) method to simultaneously position the unknown nodes and probe the abnormal nodes in complex environments. Specifically, a Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) model is constructed by introducing Mixture of Gaussian distribution and $\ell_{2,1}$-norm into the conventional low-rank Matrix Decomposition (MD) model, which can not only well fit the intrinsic low-rank property of Euclidean distance matrix, but also is robust against the node anomaly and a wider range of complex noise distributions beyond Gaussian and Laplacian noises. Our basic idea is to encode the node anomaly and the complex noise as a structural row/column sparsity matrix and a noise matrix, respectively, with entries that satisfy Mixture of Gaussian (MoG) distribution. Here, we prefer to employ MoG distribution as the general noise model due to its universal approximation property to any continuous distribution [9]. Such idea is inspired by some recent noise modeling works including Low-Rank Matrix Factorization (LRMF) [10] and Low-Rank Representation (LRR) [11], and has been verified to be effective in the complex noise scenarios. Note that although anomaly can also be seen as a continuous distribution that is approximated by MoG in theory, we still explicitly model it by using $\ell_{2,1}$-norm. The reason is to explicitly detect the location of anomaly node, and thus provide a basis for troubleshooting. Furthermore, based on the popular Expectation Maximization (EM) method [12], an efficient optimization algorithm for solving the proposed RRMD model is designed to obtain the true underlying Euclidean Distance Matrix (EDM). Finally, the actual coordinates of unknown nodes could be easily estimated by employing Multi-Dimensional Scaling (MDS) method [13], and the abnormal nodes can also be detected.

The primary contributions of our work can be summarized as follows.

1.　A Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) model is proposed to jointly estimate the missing range measurements and detect the node anomaly, which takes advantage of the potential relationship between two tasks which could help each other to achieve more accurate performance.

2.　The MoG distribution is employed to fit the unknown complex noise, which allows the proposed RRMD model to adaptively handle a wider range of noise beyond the existing methods. Meanwhile, an efficient optimization algorithm is designed to solve the proposed RRMD model by adopting the popular EM method.

3.   A novel Anomaly-aware Node Localization (ANLoC) method is proposed based on the RRMD model, and extensive experiments verify the superior positioning performance of the ANLoC method in the coexistence of node anomaly and complex noise.

The rest of this paper is organized as follows. In Section 2, we introduce the current research advances about the range-based node localization methods and low-rank matrix decomposition methods, respectively. Section 3 describes the notations used in this paper and some related mathematical foundations. In Section 4, the RRMD model is constructed and an optimization algorithm based on the EM method is designed to solve this model. This section also presents the Anomaly-aware Node Localization (ANLoC) method based on the proposed RRMD and the classic MDS. In Section 5, a series of simulation experiments are conducted to evaluate the performance of our proposed method. Finally, the conclusions are drawn in Section 6.

2. Related Work

2.1. Range-Based Node Localization

At present, the typical range-based localization methods include two steps: (1) using certain ranging methods to measure the distance between nodes, such as Time of Arrival (ToA), Time Difference of Arrival (TDoA) and Received Signal Strength Indicator (RSSI); (2) using the range measurements combined with the location of anchor nodes to calculate the position information of unknown nodes. The popular localization method called Maximum Likelihood (ML) is asymptotically efficient [14] with enough data records. Tomic et al. [15] built a new convex estimator that approximated the ML by applying efficient convex relaxations, which reduced the estimation errors. In references [16,17], WSNs localization problem is treated as a variant of EDM recovery problem or graph implementation problem. By using the range measurements between nodes and introducing the slack variable to convert non-convex quadratic distance constraints into linear constraints, the authors formulated the WSNs localization problem as a Semi-Definite Programming (SDP) problem, and designed an efficient optimization method to solve the proposed problem. References [18,19] employed the classical MDS method to map the distance relationship between wireless sensor nodes to low-dimensional space, and generated a relative coordinate map which fitted well the distance relationship between nodes. Then, a few anchor nodes were used to convert the relative position to the global position. However, the above methods did not work well in real application scenarios. In general, these methods required complete and accurate EDM. Unfortunately, due to the requirement of energy-saving and the effects of potential node anomaly and complex noise, the existing range-based node localization methods often suffer from incomplete and detorted range measurements. In addition, the ML method and SDP method could be only used to solve a small-scale problem because of its high computational complexity.

To address the above issues, some matrix completion based node localization methods have been proposed in recent years. Specifically, by taking advantage of the low-rank characteristics of EDM, Feng et al. [20] firstly formulated the range measurement imputation problem as a low-rank EDM Matrix Completion (MC) model, and designed an efficient optimization method to solve the proposed MC model. However, their work is limited to assume the errors contained in the range measurements are Gaussian noise, ignoring the existence of some complex errors caused by node hardware failure, multipath transmission, etc. In reference [6], another kind of errors called outliers was assumed to obey Laplacian distribution. Then, ℓ_1-norm regularization was introduced to deal with it, which effectively improved the positioning accuracy. However, it was also too simple to preset the complex noise to these two known types, and the actual noises should be more complicated in practical applications. Recently, Liu et al. [21] proposed a Linear Bregman Iteration based matrix completion method to localize node position for WSNs. However, this method did not consider the actual scenario under the co-existence of complex noise and anomaly. More importantly, all these matrix completion based node localization methods inevitably involved Singular Value Decomposition (SVD) operation with heavy computation cost. To address these issues, an ANLoC method is proposed in this paper, which

formulate the range measurements imputation problem as low-rank matrix decomposition instead of low-rank matrix completion model. Therefore, in the next section, we will introduce the related work on low-rank matrix decomposition.

2.2. Low-Rank Matrix Decomposition

Low-Rank Matrix Decomposition (LRMD) is an important technique in data science, which can uncover the latent manifold structures of data, and thus obtain a low dimensional compression representation. Recently, by decomposing the target matrix into the product of two low-rank matrices, LRMD has been widely used in various fields such as dimensionality reduction, clustering, and matrix recovery.

The original LRMD [22] model can be formulated as

$$\min_{U,V} \|M - UV^T\|_{\ell_p} \tag{1}$$

where $M \in \mathbb{R}^{m \times n}$ is the target matrix to be approximated, $U \in \mathbb{R}^{m \times r}$ and $V \in \mathbb{R}^{n \times r}$ are two low-dimensional matrix variables ($r < min(m, n)$), $\|\cdot\|_{\ell_p}$ denotes the ℓ_p-norm, and ℓ_1-norm and ℓ_2-norm are commonly used to make this model robust to Gaussian noise and outlier noise, respectively. When some elements of M are missing, the original LRMD could be changed into a matrix recovery model by adding an orthogonal projection operator [23], which could be formulated as:

$$\min_{U,V} \|P_\Omega(M - UV^T)\|_{\ell_p} \tag{2}$$

where $P_\Omega(\cdot)$ is an orthogonal projection operator defined as

$$[P_\Omega(M)]_{ij} = \begin{cases} M_{ij}, & (i, j) \in \Omega, \\ 0, & otherwise, \end{cases} \tag{3}$$

where $\Omega \subseteq [m] \times [n]([m] = \{1, 2, \ldots, m\}, [n] = \{1, 2, \ldots, n\})$ represents the index set of sampled elements. This model could be applied to various matrix recovery problems such as recommender system [24] and image representation [25], and has lower computational complexity than low-rank matrix completion model with nuclear norm minimization. However, ℓ_1-norm and ℓ_2-norm regularizations are only optimal when the noise follows a Gaussian or Laplacian distribution, which is not in line with the actual situation. To address this limitation, Meng et al. [10] proposed a noise-tolerant LRMD model based on Mixture of Gaussian distribution, which could be described as:

$$\min_{U,V} \|P_\Omega(M - UV^T)\|_{MoG} \tag{4}$$

where $\|\cdot\|_{MoG}$ represents that each element of this matrix is modeled by Mixture of Gaussian distribution. Reference [10] has demonstrated the robustness of the MoG model to unknown noise. However, since this model does not explicitly consider the row/column-wise structural anomaly that the sampled matrix may suffer from, it cannot be directly applied to the node localization of WSNs in this paper.

3. Preliminaries

3.1. Notations

We first introduce the important notations used in this paper. All italic letters denote variables and non-italic letters denote constants. Bold uppercase letters denote matrices and bold lowercase letters denote vectors. Specifically, X_{ij} denotes the scalar in the i-th row and j-th column of X. x^i and x_j represent the i-th row and j-th column of the matrix X respectively. Additionally, we denote the ℓ_p-norm,

Frobenius-norm, and $\ell_{2,1}$-norm of matrix $X \in \mathbb{R}^{m \times n}$ as $\left(\sum_{i=1}^{m} \sum_{j=1}^{n} |X_{ij}|^p \right)^{1/p}$, $\|X\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}^2 \right)^{1/2}$, and

$\|X\|_{2,1} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} X_{ij}^2 \right)^{1/2}$, respectively. X^T denotes the transpose of matrix X. $\mathcal{N}\left(0, \sigma^2\right)$ denotes a Gaussian distribution with mean 0 and variance σ^2. Finally, $\odot$ represents the Hadamard product of two matrices.

3.2. Mathematical Foundation

Definition 1 [26]: **(Proximal Operator)** *Let* $F(X)$ *be a real convex function defined on* $X \in \mathbb{R}^{m \times n}$, *for any* τ, $\mu > 0$ *and constant matrix* $M \in \mathbb{R}^{m \times n}$, *the proximal operator could be defined as:*

$$prox_{\tau F(X)}(M) = \arg\min_{X} \tau F(X) + \frac{\mu}{2} \|X - M\|_F^2 \tag{5}$$

Definition 2 [27]: **(Structural Thresholding Operator)** *For any* $\tau > 0, M \in \mathbb{R}^{m \times n}$, *the proximal operator of* $\ell_{2,1}$-*norm, i.e., structural thresholding operator, is defined as:*

$$prox_{\tau \|X\|_{2,1}}(M) = \mathcal{J}_{\tau/\mu}(M) \tag{6}$$

where

$$\left(\mathcal{J}_{\tau/\mu}(M) \right)^i = max\left\{ \|M^i\|_2 - \tau/\mu, 0 \right\} \cdot M^i / \|M^i\|_2, \ i = 1, 2, \ldots, m \tag{7}$$

Definition 3 [28]: **(Lipschitz Continuous Gradient)** *A differentiable convex function* $F(X)$ *defined on* $\mathbb{R}^{m \times n}$ *is said to have a Lipschitz continuous gradient* ξ, *i.e., for* $\forall X_1, X_2 \in \mathbb{R}^{m \times n}$, *there exists a constant* $\xi > 0$, *such that*

$$\left(\mathcal{J}_{\tau/\mu}(M) \right)^i = max\left\{ \|M^i\|_2 - \tau/\mu, 0 \right\} \cdot M^i / \|M^i\|_2, \ i = 1, 2, \ldots, m \tag{8}$$

Theorem 1 [29]: *Let* F_1 *and* F_2 *be two lower semi-continuous convex functions on* $X \in \mathbb{R}^{m \times n}$, *and* F_2 *is also a differentiable function with Lipschitz continuous gradient* ξ. *Then, for a convex optimization problem defined as*

$$\min_{X} F_1(X) + F_2(X) \tag{9}$$

if $F_1 + F_2$ *is mandatory and strictly convex, then for any initial value* X_0 *and* $0 < \delta < 1/\xi$, *the iterative sequence* X_{k+1} *generated by the following Equation (10) is the unique solution of problem (9)*

$$X_{k+1} = \arg\min_{X} \delta F_1(X) + \frac{1}{2} \|X - (X_k - \delta \nabla F_2(X_k))\|_F^2 \tag{10}$$

4. Anomaly-Aware Node Localization for WSNs

In this section, we first establish a RRMD model and employ EM method to optimize it, and then a novel ANLoC method is proposed based on the RRMD model.

4.1. Euclidean Distance Matrix Completion

4.1.1. Problem Description and RRMD Model Construction

In a typical application scenario, n sensor nodes are randomly deployed in a specific area, whose coordinates can be formulated as $X = [x_1, x_2, \ldots, x_n] \in \mathbb{R}^{d \times n}$ (*d* usually be 2 or 3, indicating

the dimension of the coordinate space). Then, the corresponding EDM matrix $D \in \mathbb{R}^{n \times n}$ could be calculated by

$$D_{ij} = \|x_i - x_j\|^2 = x_i^T x_i + x_j^T x_j - 2 x_i^T x_j, i, j \in \{1, 2, \dots, n\} \tag{11}$$

where D_{ij} represents the squared Euclidean distance between the i-th sensor node and the j-th sensor node. Reference [8] has proofed that the rank of matrix D is at most $d + 2$, which indicates the EDM matrix has an inherently strict low-rank property. Usually, in practical applications, only a few sensor nodes could obtain their accurate coordinates by loading GPS, and the position information of other unknown nodes need be indirectly calculated by employing some localization methods. Due to irregular node distribution and energy consumption limitation, only the partial range measurements could be collected. Based on the collected incomplete distance measurements, we can only establish a sampled EDM matrix with missing elements. Our goal is to estimate the coordinates of unknown nodes based on the sampled range measurements and the known coordinates of anchor nodes. Figure 1 illustrates the pipeline of node localization for WSNs in complex environments. As shown in Figure 1, the node localization process involves two main steps: EDM recovery and node positioning. At the first step of EDM recovery, by using our proposed RRMD algorithm, we can impute the missing range measurements and obtain the estimated complete EDM matrix, which is also called the true underlying EDM matrix, and then at the second step of node positioning, based on the estimated true underlying EDM matrix, we can further employ MDS method to position each unknown node. Compared to the second step of node positioning, the first step of EDM recovery is more critical and challenging. Specifically, in practical applications, limited by the complex environmental interference, malicious attacks, hardware malfunction, etc., the existence of errors in range measurements is unavoidable, which leads to the destruction of sampled EDM and reduction in localization accuracy. In this paper, we assume that the sampled EDM matrix $M \in \mathbb{R}^{n \times n}$ consists of the true underlying EDM component $\hat{D} \in \mathbb{R}^{n \times n}$, the row structural anomaly matrix component $R \in \mathbb{R}^{n \times n}$, the column structural anomaly matrix component $C \in \mathbb{R}^{n \times n}$, and the noise matrix component $N \in \mathbb{R}^{n \times n}$, which is illustrated in Figure 2. The noise is considered to be caused by environmental interference, malicious attacks, etc., which usually follows continuous complex distributions. Moreover, the existence of the anomaly is caused by abnormal transmission module or other unpredictable hardware defects. Specifically, when the receiving module of the nodes fails, it will cause row structural anomaly in the corresponding row in EDM, and when the sending module of nodes fails, the corresponding column in the EDM will have the column structural anomaly [8]. To address the above problems, we need to design an efficient method to obtain a true underlying EDM. Finally, based on this EDM, the actual coordinates of unknown nodes could be easily estimated by employing the classic MDS method.

Therefore, our primary goal is to design an effective EDM reconstruction method to impute the missing range measurements and de-noise the inaccurate ones. Intuitively, based on the low-rank characteristics of EDM, we can classify the EDM reconstruction problem into a standard matrix decomposition model, which is formulated as

$$\min_{U, V} \|P_\Omega(M - UV^T)\|_F. \tag{12}$$

However, this model can only be applied to the case of single Gaussian noise, and its reconstruction performance will decline dramatically when the sampled EDM is under co-existence of complex noise and anomaly. To this end, we employ MoG distribution and $\ell_{2,1}$-norm to improve the model, and then establish a Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) model as follows:

$$\min_{\hat{D}, R, C, N, U, V, \Pi, \Sigma} \lambda_1 \|R\|_{2,1} + \lambda_2 \|C^T\|_{2,1} - \log \mathcal{L}(P_\Omega(N) | \Pi, \Sigma),$$

$$s.t. \ P_\Omega(M) = P_\Omega(\hat{D} + N + R + C), \hat{D} = UV^T, \sum_{k=1}^{n_c} \pi_k = 1, \pi_k \geq 0, \ k = 1, 2, \cdots, n_c, \tag{13}$$

where $M \in \mathbb{R}^{n \times n}$ denotes the sampled matrix, $R \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{n \times n}$ represent row structural anomaly and column structural anomaly, respectively, $U, V \in \mathbb{R}^{n \times (d+2)}$ are two low-rank components of the true underlying EDM matrix $\hat{D}$, $log\mathcal{L}(P_{\Omega}(N)|\Pi, \Sigma)$ is the log-likelihood function, $\Pi = \{\pi_1, \pi_{2,...,} \pi_{n_c}\}, \Sigma = \left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_c}^2\right\}$ and n_c are the parameters of MoG, and λ_1 and λ_2 are the tunable parameters. The key of this model is that we use MoG distribution to smooth any unknown noises and $\ell_{2,1}$-norm to detect node anomaly, respectively. The main motivations for this model are as follows. (1) First of all, in response to the complex noise, we introduce a Mixture of Gaussians distribution, which is treated as a universal approximator to any continuous density function. Therefore, the RRMD model is robust against a wider range of complex noise distributions beyond Gaussian and Laplacian noises. (2) Then, as the optimal measure of row sparsity matrix, $\ell_{2,1}$-norm could be used to smooth row-wise anomaly and the transpose of column-wise anomaly, thus the abnormal nodes can be detected. (3) Finally, in order to deal with the challenge of data missing, we choose the classic low-rank MD method, which can impute the missing elements with lower computational costs than the low-rank MC method.
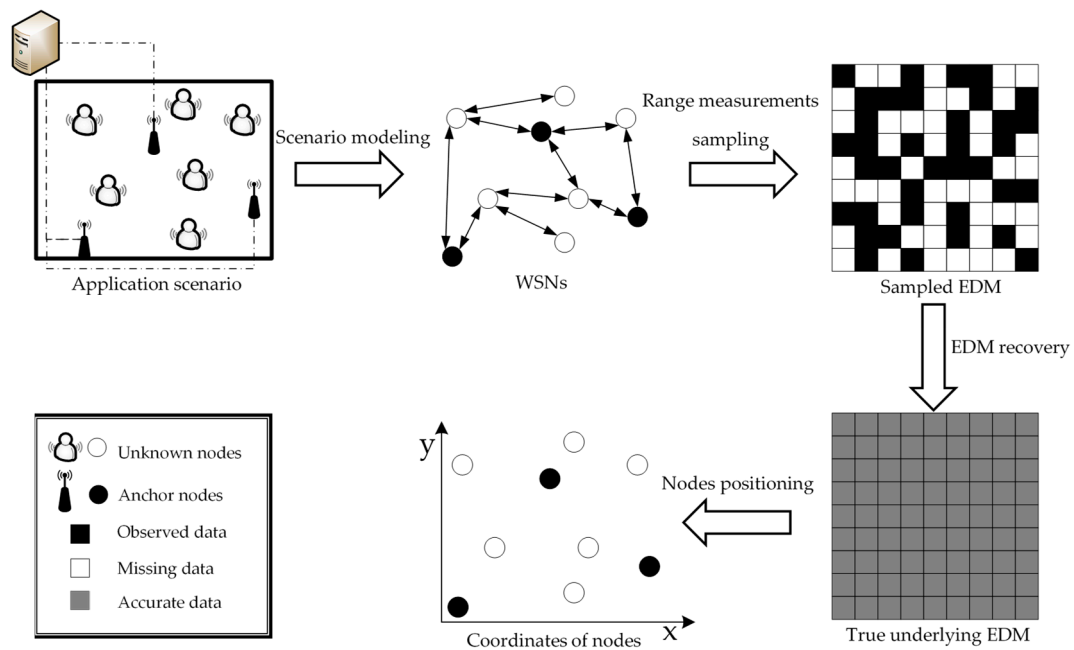

Figure 1. Pipeline of node localization for Wireless Sensor Networks (WSNs) in complex environments.
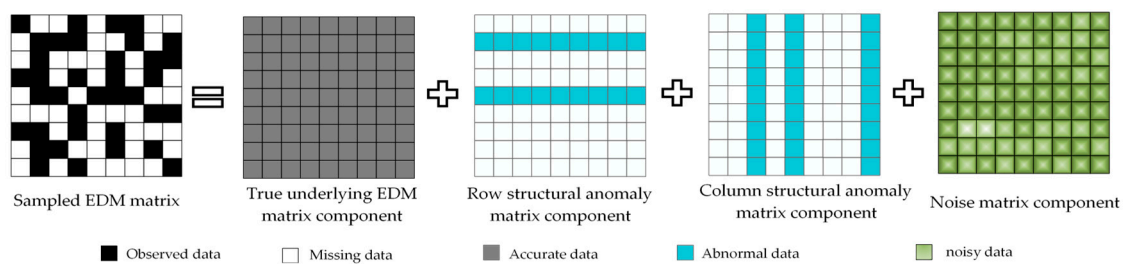

Figure 2. Illustration of the sampled Euclidean Distance Matrix (EDM) matrix.

In this paragraph, we introduce how to apply MoG distribution to the RRMD model. Without loss of generality, each element $N_{ij}(i, j = 1, 2, \ldots, n)$ in the noise matrix N is considered to be from a Mixture of Gaussians distribution, which is defined as

$$\mathbb{P}\left(N_{ij}\right) \sim \sum_{k=1}^{n_c} \pi_k \mathcal{N}\left(N_{ij}|0, \sigma_k^2\right) \tag{14}$$

where $\mathcal{N}\left(N_{ij}|0, \sigma_k^2\right)$ denotes the Gaussian distribution with mean 0 and variance σ^2. n_c is the number of Gaussian components and $\pi_k \geq 0$ represents the mixing proportion where $\sum_{k=1}^{n_c} \pi_k = 1$. Then, the likelihood function can be written as

$$\mathcal{L}(P_\Omega(N)|\Pi, \Sigma) = \prod_{i,j \in \Omega} \sum_{k=1}^{n_c} \pi_k \mathcal{N}\left(N_{ij}, 0, \sigma_k^2\right) \tag{15}$$

where $\Pi = \{\pi_1, \pi_{2,\ldots,} \pi_{n_c}\}$, $\Sigma = \left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_c}^2\right\}$. Usually, we use the log-likelihood function instead of this likelihood function for convenient calculation. Then, our aim is to maximize the log-likelihood function to obtain the MoG parameters. Obviously, we can construct the final objective function expressed as Equation (13) by combining the negative log-likelihood function with the $\ell_{2,1}$-norm, and then minimize it to get all unknown variables in RRMD model.

4.1.2. Optimizing RRMD via Expectation Maximization Method

In this section, we employ the popular EM method to solve the proposed RRMD model. Firstly, we explain the motivation for using this method. Obviously, in order to gain the estimated values of parameters Π, Σ, the log-likelihood function $log\mathcal{L}(P_\Omega(N)|\Pi, \Sigma)$ should be maximized. However, for many specific problems, parameters Π, Σ cannot be directly calculated due to complex expressions of log-likelihood functions. Fortunately, the EM method can be employed to solve such problems, which has been proven to be effective by many researches. Specifically, let $Z_{ijk} \in \{0, 1\}$ be a set of hidden variables with $\sum_{k=1}^{n_c} Z_{ijk} = 1(i, j = 1, 2, \ldots, n)$, and $Z_{ijk} = 1$ implies the noise N_{ij} comes from the k-th Gaussian component. Therefore, the log-likelihood function can be re-written in a form that is easy to optimize as

$$log\mathcal{L}(P_\Omega(N)|\Pi, \Sigma) = \sum_{i,j \in \Omega} \sum_{k=1}^{n_c} Z_{ijk}\left(log\pi_k - log\sqrt{2\pi}\sigma_k - \frac{1}{2\sigma_k^2}\left(M_{ij} - u_i^T v_j - R_{ij} - C_{ij}\right)^2\right) \tag{16}$$

Next, we will introduce the specific optimization process of estimating the parameters $(\hat{D}, R, C, N, U, V, \Pi, \Sigma)$ in Equation (13) by alternately conducting the E and M steps.

- E step:

In the E step, we calculate the conditional expectation Y_{ijk} of Z_{ijk}. The specific calculation formula can be given as follows.

$$Y_{ijk} = E\left(Z_{ijk}\right) = \frac{\pi_k \mathcal{N}\left(M_{ij} - u_i^T v_j - R_{ij} - C_{ij}, 0, \sigma_k^2\right)}{\sum_{k=1}^{n_c} \pi_k N\left(M_{ij} - u_i^T v_j - R_{ij} - C_{ij}, 0, \sigma_k^2\right)} \tag{17}$$

- M step:

In the M step, we need to optimize the following optimization model as

$$\min_{U,V,R,C,\Pi,\Sigma} \lambda_1 \|R\|_{2,1} + \lambda_2 \|C^T\|_{2,1}$$
$$-\sum_{i,j \in \Omega} \sum_{k=1}^{n_c} Y_{ijk}\left(log\pi_k - log\sqrt{2\pi}\sigma_k - \frac{1}{2\sigma_k^2}\left(M_{ij} - u_i^T v_j - R_{ij} - C_{ij}\right)^2\right) \tag{18}$$
$$s.\,t.\ \sum_{k=1}^{n_c} \pi_k = 1, \pi_k \geq 0$$

A natural idea is to use alternate iteration methods to solve MoG parameters Π, Σ and matrix variables $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{R}, \boldsymbol{C}$, i.e.,

Update Π, Σ:

$$\pi_k = B_k / \sum_{k=1}^{n_c} B_k, \text{ where } B_k = \sum_{i,j \in \Omega} Y_{ijk} \tag{19}$$

$$\sigma_k^2 = \frac{1}{B_k} \sum_{i,j \in \Omega} Y_{ijk} \left(M_{ij} - \boldsymbol{u}_i^T \boldsymbol{v}_j - R_{ij} - C_{ij} \right)^2 \tag{20}$$

Update $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{R}, \boldsymbol{C}$:

In order to update $\boldsymbol{U}, \boldsymbol{V}, \boldsymbol{R}, \boldsymbol{C}$, we need to solve the following sub-problem:

$$\min_{\boldsymbol{U},\boldsymbol{V},\boldsymbol{R},\boldsymbol{C}} \lambda_1 \|\boldsymbol{R}\|_{2,1} + \lambda_2 \|\boldsymbol{C}^T\|_{2,1} + \sum_{i,j \in \Omega} \sum_{k=1}^{n_c} Y_{ijk} \frac{\left(M_{ij} - \boldsymbol{u}_i^T \boldsymbol{v}_j - R_{ij} - C_{ij} \right)^2}{2\sigma_k^2} \tag{21}$$

If we let

$$W_{ij} = \begin{cases} \sqrt{\sum_{k=1}^{n_c} Y_{ijk}/\sigma_k^2}, & (i,j) \in \Omega, \\ 0 & , \quad otherwise. \end{cases} \tag{22}$$

Then Equation (21) can be reformulated as follows:

$$\min_{\boldsymbol{U},\boldsymbol{V},\boldsymbol{R},\boldsymbol{C}} \lambda_1 \|\boldsymbol{R}\|_{2,1} + \lambda_2 \|\boldsymbol{C}^T\|_{2,1} + \frac{1}{2} \|\boldsymbol{W} \odot \left(\boldsymbol{M} - \boldsymbol{U}\boldsymbol{V}^T - \boldsymbol{R} - \boldsymbol{C} \right)\|_F^2 \tag{23}$$

Furthermore, we can optimize Equation (23) by solving the following three sub-problems.

1) Update $\boldsymbol{U}, \boldsymbol{V}$ by solve the following sub-problem:

$$\min_{\boldsymbol{U},\boldsymbol{V}} \frac{1}{2} \|\boldsymbol{W} \odot \left(\boldsymbol{M} - \boldsymbol{U}\boldsymbol{V}^T - \boldsymbol{R} - \boldsymbol{C} \right)\|_F^2 \tag{24}$$

Obviously, as a typical weighted matrix decomposition problem, it could be solved by various methods, such as ALS [30], WLRA [31] and DN [32].

2) Update R by conducting the following iteration:

$$\boldsymbol{R}_t = \mathcal{J}_{\lambda_1 \tau_R} \left(\boldsymbol{R}_{t-1} - \tau_R \boldsymbol{W} \odot \boldsymbol{W} \odot \left(\boldsymbol{R}_{t-1} + \boldsymbol{C} + \boldsymbol{U}\boldsymbol{V}^T - \boldsymbol{M} \right) \right) \tag{25}$$

where t represents the t-th iteration and τ_R is the step size of the proximal gradient.

3) Update $\boldsymbol{C}$ by conducting the following iteration:

$$\left(\boldsymbol{C}^T \right)_t = \mathcal{J}_{\lambda_2 \tau_C} \left(\left(\boldsymbol{C}_{t-1} - \tau_c \boldsymbol{W} \odot \boldsymbol{W} \odot \left(\boldsymbol{R} + \boldsymbol{C}_{t-1} + \boldsymbol{U}\boldsymbol{V}^T - \boldsymbol{M} \right) \right)^T \right) \tag{26}$$

where t represents the t-th iteration and τ_C is the step size of the proximal gradient.

Based on the aforementioned analysis, we summarize the whole procedure in Algorithm 1.

Algorithm 1. Proposed Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) Algorithm

---

Input: The sampled EDM matrix M, the index set Ω, the parameters λ_1, λ_2 and threshold θ, the initial number of Gaussian components n_c.

Output: The true underlying EDM matrix $\hat{D}$, the row structural anomaly R, and the column structural anomaly C.

1. Randomly initialize $U_0, V_0, \Pi_0, \Sigma_0$; initialize R_0, C_0 to zero matrix;
2. $t = 1$;
3. While not convergence do
4.   (E-step)**:** update Υ according to Equation (17);
5.   (M-step)**:** update Π, Σ according to Equation (19) and Equation (20);
6.   (M-step): update U, V according to optimize Equation (24);
7.   (M-step): update R according to Equation (25);
8.   (M-step): update C according to Equation (26);
9.   (Tuning n_c): Let g_i and g_i represent the number of i-th and j-th Gaussian component respectively. if $|\sigma_i^2 - \sigma_j^2|/(\sigma_i^2 + \sigma_j^2) < \theta$, then let $\pi_i = \pi_i + \pi_j$, $\sigma_i^2 = (g_i\sigma_i^2 + g_j\sigma_j^2)/(g_i + g_j)$, $K = K - 1$. Lastly, remove π_j and σ_j^2 from Π, Σ, respectively.
10. $t = t + 1$;
11. End while
12. R, C and $\hat{D} = UV^T$.

4.2. Anomaly-Aware Node Localization

Based on the proposed RRMD model, all pair-wise range measurements between nodes can be easily obtained. However, it is only the first step of this range-based localization method. The second step of this method is to calculate the location information of the unknown nodes based on the complete distance information and the actual coordinates of anchor nodes, and which can be implemented by using the classical MDS method [6,8,21].

MDS is a typical low-dimensional embedding method, which is originally proposed to solve the curse of dimensionality in the field of machine learning. The MDS method ensures that the distance between samples in the original high-dimensional space can be preserved in low-dimensional space. The input of this method can be high-dimensional features of samples (the pair-wise distance information can be easily obtained from high-dimensional features) or the distance information between samples, while the output is the low-dimensional features of samples in the specified d-dimensional space. Therefore, in the sensor networks localization application, we can employ the MDS method to calculate the relative coordinates of each node in the d-dimensional space. Then, based on the actual/absolute coordinates of at least $d + 1$ anchor nodes and the relative coordinates obtained, the coordinate transformation matrix between the actual coordinates and the relative coordinates can be calculated (with the details provided in Theorem 2). Finally, the relative coordinates can be mapped to the actual ones by using the coordinate transformation matrix.

Next, we introduce the specific steps to implement the above MDS method. Firstly, we let the relative coordinates and actual coordinates of n nodes be represented as $T = [t_1, t_2, \ldots, t_n] \in \mathbb{R}^{d \times n}$, $A = [a_1, a_2, \ldots, a_n] \in \mathbb{R}^{d \times n}$, respectively, and assume that nodes $1, 2, \ldots n_a (n_a \geq d + 1)$ are anchor nodes. Secondly, the classic MDS method is employed to calculate the relative coordinates of all nodes. Thirdly, with the actual coordinates of anchor nodes, we have coordinate transformation matrix:

$$Q = \frac{[a_2 - a_1, a_3 - a_1, \ldots, a_{n_a} - a_1]}{[t_2 - t_1, t_3 - t_1, \ldots, t_{n_a} - t_1]} \tag{27}$$

Finally, the coordinate of all unknown nodes can be obtained by

$$(a_i = Q \cdot (t_i - t_1) + a_1, i = n_a + 1, n_a + 2, \ldots n \tag{28}$$

Based on the aforementioned analysis, the proposed Anomaly-aware Node Localization (ANLoC) algorithm can be summarized as in Algorithm 2.

Algorithm 2. Anomaly-aware Node Localization (ANLoC) Algorithm

---

Input: The sampled EDM matrix M, the index set Ω, the parameters λ_1, λ_2 and threshold θ, the initial number of Gaussian components n_c. The coordinates of anchor nodes $\{a_1, a_2, \ldots, a_{n_a}\}(n_a \geq d+1)$, where n_a denotes the number of anchor nodes.

Output: Coordinates of all unknown nodes $\{a_i | i = n_a + 1, n_a + 2, \ldots n\}$.

1. Calculate the true underlying EDM matrix $\hat{D}$ by using Algorithm 1;
2. Double centering the matrix $\hat{D}$:
     $S = -0.5 \times J\hat{D}J$, where $J = I - 1 \cdot 1^T / n$ and I is identity matrix;
3. Perform SVD decomposition on matrix S:
$$[H, \Lambda, K] = svd(S);$$
4. Calculate relative coordinates:
$$T = [t_1, t_2, \ldots, t_n] = \sqrt{\Lambda_d} \cdot H_d^T$$
     where $t_i \in \mathbb{R}^{d \times 1}$, $\Lambda_d = \Lambda(1:d, 1:d)$, $H_d = H(:, 1:d)$;
5. Calculate the coordinate transformation matrix Q:
     $Q = [a_2 - a_1, a_3 - a_1, \ldots, a_{n_a} - a_1] / [t_2 - t_1, t_3 - t_1, \ldots, t_{n_a} - t_1]$;
6. Calculate and output coordinates of all unknown nodes:
$$a_i = Q \cdot (t_i - t_1) + a_1, i = n_a + 1, n_a + 2, \ldots n.$$

Theorem 2: *For sensor nodes localization in d-dimensional space, given the absolute positions $\{a_1, a_2, \cdots, a_{n_a}\}$ of the n_a anchor nodes and the relative coordinates $\{t_1, t_2, \cdots, t_n\}$ of all the n $(n \gg n_a)$ sensor nodes, if $n_a \geq d + 1$, then the relative coordinates $\{t_{n_a+1}, t_{n_a+2}, \cdots, t_n\}$ can be transformed to the corresponding absolute positions $\{a_{n_a+1}, a_{n_a+2}, \cdots, a_n\}$.*

Proof: According to reference [13], the absolute positions a_i $(i \geq n_a + 1)$ can be computed according to

$$a_i = Q \cdot (t_i - t_1) + a_1, i = n_a + 1, n_a + 2, \cdots, n \tag{29}$$

where $Q \in \mathbb{R}^{d \times d}$ is the unknown coordinate-transform matrix, which should be determined by the following matrix equation:

$$Q \cdot [t_2 - t_1, t_3 - t_1, \cdots, t_{n_a} - t_1] = [a_2 - a_1, a_3 - a_1, \cdots, a_{n_a} - a_1] \tag{30}$$

Without loss of generality, let $\hat{T} = [t_2 - t_1, t_3 - t_1, \cdots, t_{n_a} - t_1]$ and $\hat{A} = [a_2 - a_1, a_3 - a_1, \cdots, a_{n_a} - a_1]$, then we can see that $\hat{T} \in \mathbb{R}^{d \times (n_a - 1)}$ and $\hat{A} \in \mathbb{R}^{d \times (n_a - 1)}$. Therefore, the above matrix Equation can be regarded as the following equivalent equations:

$$\sum_{l=1}^{d} Q_{il} \cdot \hat{T}_{lj} = \hat{A}_{ij}, i = 1, 2, \cdots, d, \text{ and } j = 1, 2, \cdots, n_a - 1 \tag{31}$$

Obviously, if $n_a \leq d$, then the number of unknown entries (i.e., d^2) is greater than the number of Equations (i.e., $d \times (n_a - 1)$), so we cannot obtain the determined coordinate-transform matrix Q. On the contrary, if $n_a \geq d + 1$, we only need to arbitrarily select $d + 1$ anchor nodes, then we can obtain the unique solution of Q. □

5. Performance Evaluation

In this section, we first describe the experimental setting and evaluation metrics, and then report some extensive experimental results under the different scenarios.

5.1. Experimental Setting

In order to investigate the performance of our proposed ANLoC, some simulation experiments were conducted. All these experiments were designed based on MATLAB 2017a and run on PC with Intel i5-8400 CPU and 16G RAM. Firstly, we randomly disposed 100 sensor nodes in a square area with 100×100 unit (where the unit can be determined according to actual communication condition, such as meter, decimeter, foot, and inch, etc.), and 6 of them were anchor nodes while others were unknown nodes. Let $X \in \mathbb{R}^{2 \times 100}$ be the actual coordinate matrix of all nodes and $D \in \mathbb{R}^{100 \times 100}$ denote the ground truth EDM matrix between nodes. Secondly, we artificially added some error matrices into matrix D to simulate the complex noises and the structural anomaly, and thus obtained a corrupted matrix D_{error}. Specifically, the complex noise was set to the mixture of the following components: (1) Gaussian noise with mean 0 and variance 100; (2) Gaussian noise with mean 0 and variance 50; and (3) sparse noise with a pollution ratio 1% randomly generated within the range of $[0, 10000]$. Moreover, the structural anomalies were set to be randomly generated within the range of $[0, 500]$ and with a row-wise pollution ratio 3% and a column-wise pollution ratio 3%. Thirdly, sampled matrix M could be obtained by randomly sampling some elements from D_{error}. Finally, the true underlying EDM matrix can be reconstructed from M via Algorithm 1, and then the actual coordinates of all nodes could be calculated using Algorithm 2. Moreover, we summarized all the simulation parameters in Table 1.

Table 1. Simulation parameters.

Name	Value	Name	Value
Size of experimental scenario	100×100	Gaussian noise 1	Mean 0, variance 100
Number of sensor nodes	100	Gaussian noise 2	Mean 0, variance 50
Number of anchors	6	Range of sparse noise	$[0, 10000]$
Range of row anomaly	$[0, 500]$	Range of column anomaly	$[0, 500]$

The proposed ANLoC method involves three hyper-parameters λ_1, λ_2 and n_c. It is important to adjust these parameters to proper values. In order to determine the optimal value of these parameters, we employed a 10-fold cross validation procedure and conduct the grid search method. In particular, firstly, let $n_c = 5$, we searched for the optimal solution of λ_1 and λ_2 within the range of {0.001, 0.01, 0.1, 1, 10, 100, 1000} by minimizing the matrix recovery error of EDM. Then, based on the selected optimal parameter λ_1 and λ_2, n_c was determined within the range of {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}. Last but not least, we compared our method with three different competing methods, including NLIRM method [6], SVT-based method [20], and OptSpace-based method [33]. All the involved parameters in these competing methods were optimized by using the same nested 10-fold cross-validation procedure as in our ANLoC method.

5.2. Evaluation Metrics

We selected the following four evaluation metrics to evaluate the node localization performance:

- EDM recovery error:
$$e_r = \|\hat{D} - D\|_F / \|D\|_F \tag{32}$$
where $\hat{D}$ denotes the reconstructed matrix and D is the ground truth matrix.
- Localization error:
$$e_l = \|\hat{X} - X\|_F / n \tag{33}$$
where $\hat{X}$ denotes the coordinates calculated by Algorithm 2.
- Anomaly recognition accuracy:
$$p_{ave} = (p_{row} + p_{col}) / 2 \tag{34}$$

where p_{row} and p_{col} represent the row anomaly recognition accuracy and column anomaly recognition accuracy, respectively. Specifically, the p_{row} is calculated by

$$p_{row} = 2 \times \frac{r_{pre} \cdot r_{rec}}{r_{pre} + r_{rec}}, r_{pre} = \frac{r_{tru}}{r_{all}}, r_{rec} = \frac{r_{tru}}{r_{act}} \tag{35}$$

where r_{all} represents the number of abnormal rows identified by the method proposed in this paper, r_{tru} indicates the number of correctly recognized rows in r_{all} and r_{act} represents the number of actual abnormal rows; the p_{row} is calculated by

$$p_{col} = 2 \times \frac{c_{pre} \cdot r_{rec}}{c_{pre} + r_{rec}}, c_{pre} = \frac{c_{tru}}{c_{all}}, c_{rec} = \frac{c_{tru}}{c_{act}} \tag{36}$$

where c_{all} represents the number of abnormal columns identified by the method proposed in this paper, c_{tru} indicates the number of correctly recognized columns in c_{all}, and c_{act} represents the number of actual abnormal columns.

- Cumulative distribution of localization errors [21]:

$$e_{CD} = P(\Delta l_i \leq \sigma) \tag{37}$$

where $\Delta l_i = \sqrt{(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2}$ defines the localization error of *i*-th node, (x_i, y_i) and $(\hat{x}_i, \hat{y}_i)$ denote the actual coordinates and the estimated coordinates, respectively.

5.3. Experimental Results

In order to investigate the performance of proposed method in complex environments, we conducted the following four experiments under the different application scenarios.

- Scenario 1: Localization without the noise and anomaly

This experiment assumed that all observed range measurements were accurate. Figure 3 shows that the variation of the EDM recovery error and average localization error of WSNs nodes when the EDM is sampled at different proportions. The horizontal axis of the two subgraphs represents the sampling ratio, while the vertical axis in Figure 3a represents the EDM recovery error, and the vertical axis in Figure 3b represents the localization error of WSNs nodes. Comparing Figure 3a with Figure 3b, it is obvious to find that the EDM recovery error and localization error are consistent with the trend of the sampling rate. Under this scenario, when sampling rate reaches 0.3, the EDM recovery error and the localization error of these four methods reach a very low level. This shows that the existing methods and the method proposed in this paper can achieve good positioning accuracy under the ideal condition.
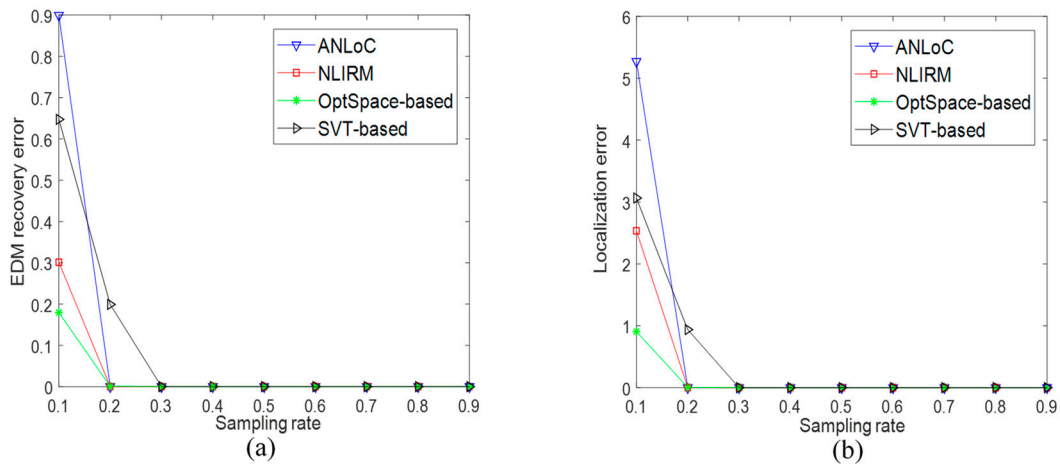
Figure 3. Performance comparison under scenario 1 without the noise and anomaly. (**a**) EDM recovery error and (**b**) localization error.

- Scenario 2: Localization with only complex noise

This experiment aims to examine the performance of four methods with only complex noise. Therefore, we only add complex noise into EDM. After sampling the EDM in different proportions, the experimental results are shown in Figure 4. Under this scenario, the SVT-based method and the OptSpace-based method have large EDM recovery error and localization error even if the sampling rate is high, which indicates that the two methods do not handle the complex noise well. When the sampling ratio reaches 0.3, the EDM recovery error of NLIRM could be stabilized at around 0.01, and the localization error of nodes could be less than 0.03. However, the EDM recovery error and localization error of ANLoC could reach 0.001 and 0.01, respectively. Comparing with Figure 3, we find that both the SVT-based method and the OptSpace-based method could not deal with complex noise at all. The NLIRM has a certain ability to resist noise, but it is slightly worse than our ANLoC. Therefore, our ANLoC could perform best under complex noise than others.

Figure 4. Performance comparison under scenario 2 with only complex noise. (**a**) EDM recovery error and (**b**) localization error.

- Scenario 3: Localization with only anomaly

The purpose of this experiment is to examine the performance of detecting anomalous nodes in the environment without any noise. In this case, only anomaly is added into EDM. As shown in Figure 5, compared with Figure 3, the EDM recovery error and the localization error of the four

methods are all increased when the EDM is destroyed by the anomaly. However, compared with other three methods, ANLoC could have better performance at a lower sampling rate (0.2). In addition, the ANLoC could also detect abnormal nodes. Specifically, Figure 5c shows that when the sampling rate reaches 0.2, the recognition accuracy of abnormal nodes can reach 100%. In comparison with Figure 3, we find that the performance of all methods has been degraded in varying degrees, but our ANLoC still has the best performance. Therefore, when some sensor nodes have abnormities, the ANLoC has good performance for precisely locating unknown nodes and detecting abnormal nodes.

Figure 5. Performance comparison under scenario 3 with only anomaly. (**a**) EDM recovery error; (**b**) localization error; and (**c**) anomaly recognition accuracy.

- Scenario 4: Localization with both complex noise and anomaly

This experiment is designed to evaluate the performance of the ANLoC in a complex environment where complex noise and anomaly coexist. The complex noise and the anomaly are both added to the EDM to simulate the impact of the complex real application scenarios. Experimental results are shown in Figure 6. Obviously, compared with the other three methods, our ANLoC method proposed in this paper not only has the lowest EDM recovery error and localization error, but also has high recognition accuracy of abnormal nodes when the sampling rate reaches 0.3. Comparing with Figure 3, we can find that the performance of the SVT-based method and the OptSpace-based method are much worser in complex environment. The NLIRM has less performance degradation, but its EDM recovery accuracy and localization accuracy are still not as high as our ANLoC. In short, our proposed ANLoC has good robustness to complex application scenarios and could detect faulty nodes well. In addition, Figure 7a shows the cumulative distribution of the localization error for the four methods when the sampling rate is 0.5 under this scenario. The probability of the localization error of the ANLoC being less than 0.5 is 98% and the probability of being less than 1 is 100% while the other three methods are all below 30%. In addition, Figure 7b shows the intuitionistic localization results of all nodes in this case when the sampling rate is 0.5.
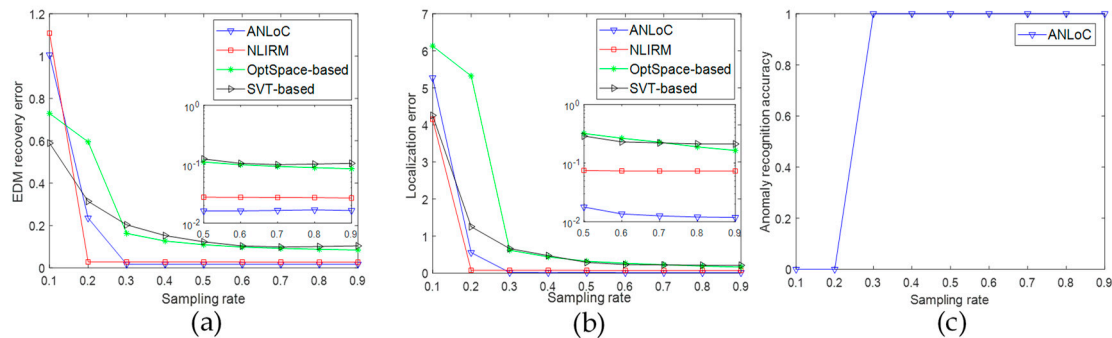

Figure 6. Performance comparison under scenario 4 with both complex noise and anomaly. (**a**) EDM recovery error; (**b**) localization error; and (**c**) anomaly recognition accuracy.
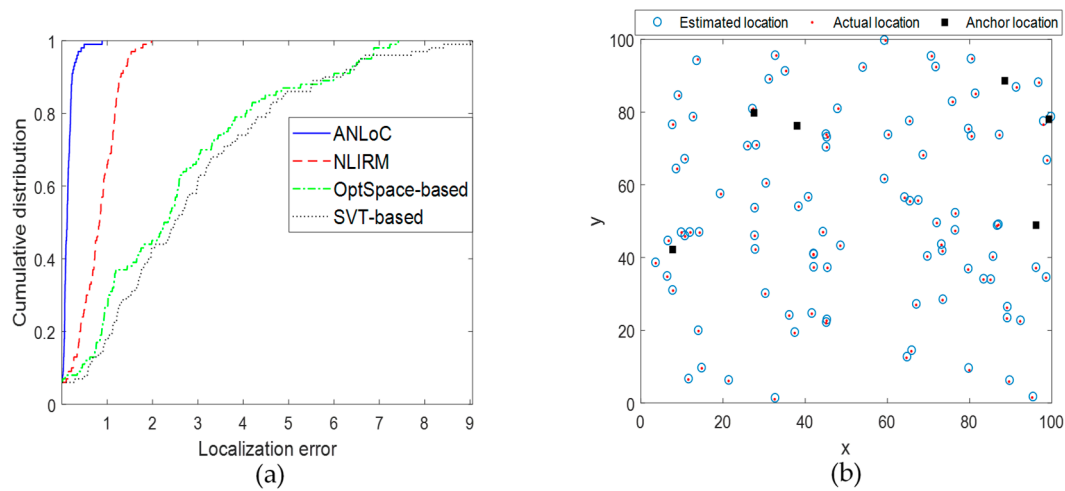
Figure 7. Illustration of localization performance under scenario 4 with both complex noise and anomaly. (**a**) Cumulative distribution of localization error and (**b**) positioning result display.

5.4. Effects of the Proposed Strategies

To analyze the effects of our proposed strategies, experiments are designed to compare the proposed method with three partial deleted methods. Actually, our RRMD can be treated as LRMD + MoG + $\ell_{2,1}$-norm. Then, the remaining three types of partial deleted methods are as follows. (1) Pure LRMD: do not consider any noise or anomaly; (2) LRMD + MoG: only consider complex noise; (3) LRMD + $\ell_{2,1}$-norm: only consider the anomaly. What's more, these experiments are set under the scenario with both complex noise and anomaly. Finally, all experimental results are shown in Figure 8a–c. Obviously, our ANLoC considers both complex noise and anomaly, which leads to its low EDM recovery error and localization error. Moreover, when the sampling rate reaches 0.3, it could accurately detect abnormal nodes. Due to too ideal assumptions, pure LRMD is the worst performer, which means that it cannot be applied in actual situation. Although LRMD + MoG could handle the error of EDM well, it does not have the ability to detect abnormal nodes. Specially, LRMD + $\ell_{2,1}$-norm could hardly detect abnormal nodes in this case. Naturally, we speculate the noise that is not unprocessed may have a negative effect on the performance for detecting abnormal nodes in this method. Therefore, another experiment is designed with LRMD + $\ell_{2,1}$-norm when the EDM is only destroyed by anomaly, and the result of anomaly recognition is drawn in Figure 8d. The pleasure is that this supplementary experiment validates our conjecture. In conclusion, these experiments demonstrate the effectiveness of the MoG and $\ell_{2,1}$-norm strategies.
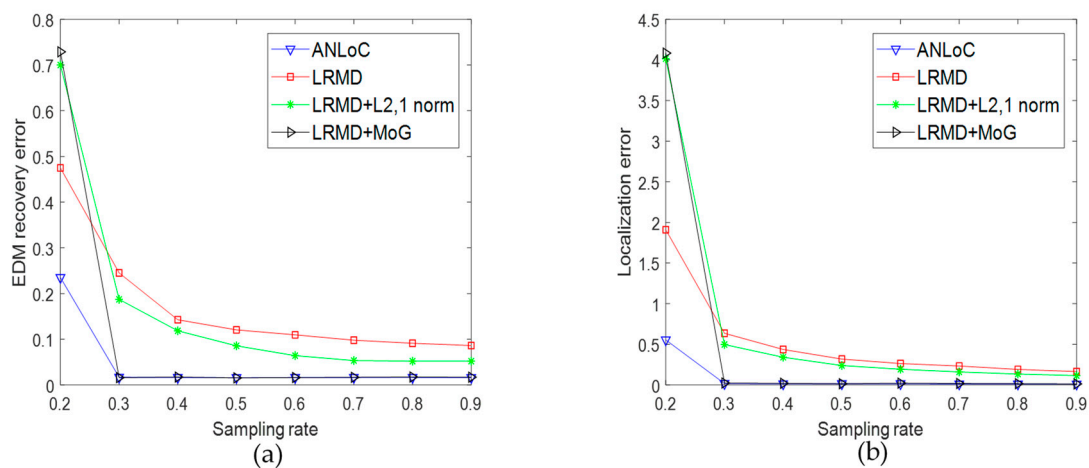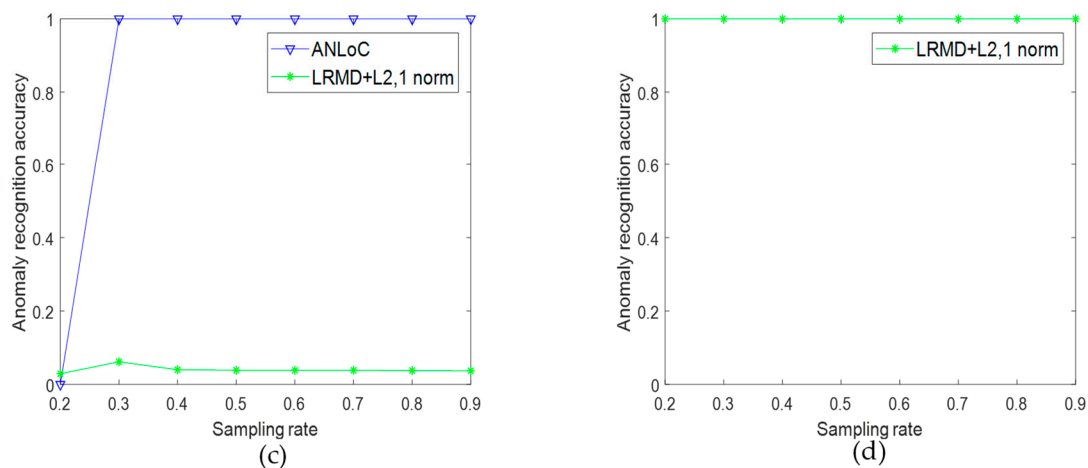

Figure 8. *Cont.*

Figure 8. Effects illustration of the proposed strategies. (**a**) EDM recovery error under the scenario 4 with both complex noise and anomaly; (**b**) localization error under the scenario 4 with complex noise and anomaly; (**c**) anomaly recognition accuracy under condition with complex noise and anomaly; and (**d**) anomaly recognition accuracy under the scenario 3 with only anomaly.

5.5. Localization for Large-Scale Scenario

In a practical application scenario, there is a case that when the size of the localization scenario is much larger than ranging length of the sensor nodes and only a few range measurements between inter-nodes could be obtained. As a result of that, the sampled EDM based on range measurements will be very sparse, which leads to poor performance of our method. To solve this problem, a large-scale localization method is proposed in this paper. For the convenience of discussion, we simplify the application scenario as shown in Figure 9. Suppose there is a 180×100 unit area divided into three parts I, II, and III from left to right, and among them, there are 80, 20, 80 nodes, respectively. Of the 180 nodes, there are 6 anchor nodes and 6 relay nodes. Then, regions I and II are combined into a localization sub-region $\mathcal{M}$ while regions II and III are combined into a localization sub-region $\mathcal{N}$. The large-scale process is as follows. (1) Calculate the actual coordinates of all nodes within sub-region $\mathcal{M}$ by ANLoC. (2) Based on step one, the actual coordinates of relay nodes are gained. Taking them as anchor nodes, we employ ANLoC again to get the coordinates of the remaining unknown nodes in region $\mathcal{N}$.
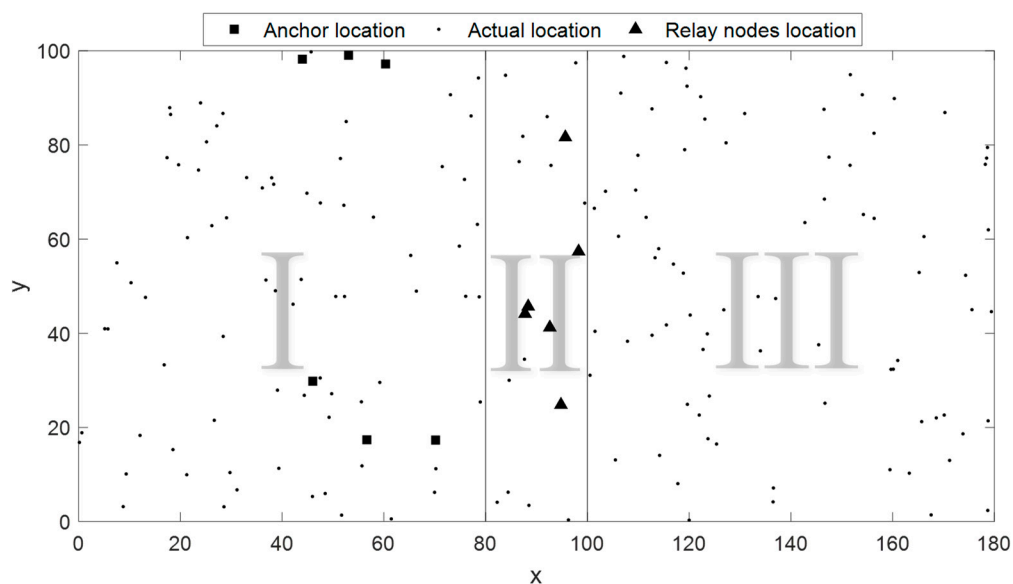

Figure 9. Illustration of large-scale scenario used in this study.

Further, in order to verify the feasibility of the above-mentioned large-scale scenario localization method, we designed an experiment with complex noise and anomaly. Moreover, the EDM sampling rates of region $\mathcal{M}$ and $\mathcal{N}$ are both set 0.5. As Figure 10 shows, the experimental results confirm that our large-scale localization method can be well applied to large-scale scenarios.
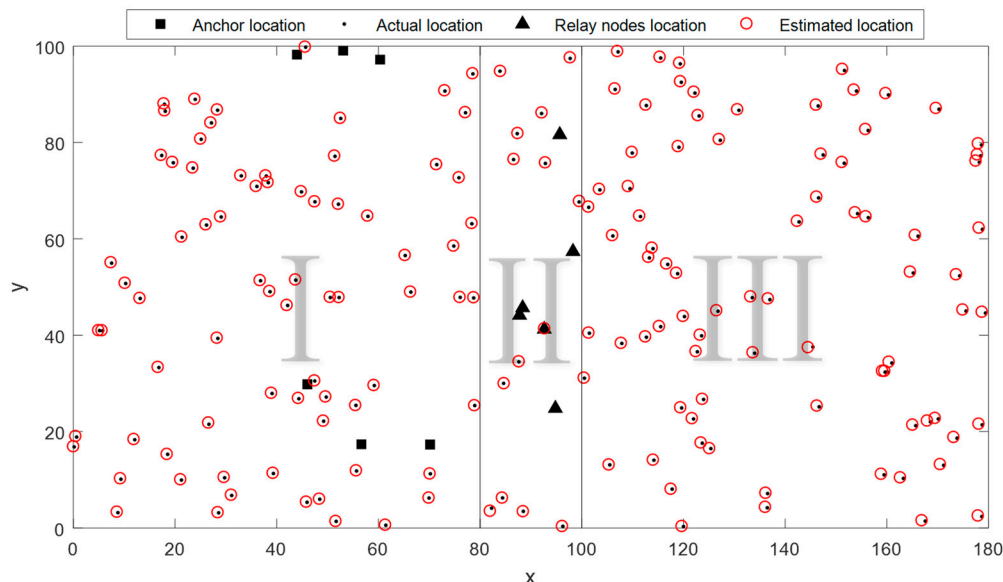

Figure 10. Localization results under large-scale scene scenario with complex noise and anomaly.

6. Conclusions

In this paper, we aim to estimate node position of WSNs in complex environments. Considering that the coexistence of potential node anomaly and complex noise in practical applications, we propose a novel Anomaly-aware Node Localization (ANLoC) method to address this task. Specifically, the proposed ANLoC method is divided into two steps. First, a Robust $\ell_{2,1}$-norm Regularized Matrix Decomposition (RRMD) model is designed to simultaneously detect node anomaly and impute the missing range measurements. Second, based on the imputed EDM matrix, all unknown nodes can be easily estimated by using the classic MDS method. The extensive experiments demonstrated our proposed ANLoC method consistently outperforms three state-of-the-art localization methods in terms of EDM recovery error, localization error, and anomaly recognition accuracy. Additionally, our proposed ANLoC method also can extend to a large-scale localization scenario. Future work will focus on extending our ANLoC approach to an incremental version, making it suitable for handling dynamic positioning of mobile sensor nodes.

Author Contributions: Conceptualization, P.X and L.C; formal analysis, P.X. and L.C.; funding acquisition, L.C.; software, P.X. and T.C.; writing—original draft, P.X.; writing—review and editing, P.X., L.C., and T.C.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zheng, K.; Wang, H.; Li, H.; Xiang, W.; Lei, L.; Qiao, J.; Shen, X.S. Energy-efficient localization and tracking of mobile devices in wireless sensor networks. *IEEE Trans. Veh. Technol.* **2017**, *66*, 2714–2726. [CrossRef]
2. Qian, H.; Fu, P.; Li, B.; Liu, J.; Yuan, X. A Novel Loss Recovery and Tracking Scheme for Maneuvering Target in Hybrid WSNs. *Sensors* **2018**, *18*, 341. [CrossRef] [PubMed]

3. Zhu, H.; Xiao, F.; Sun, L.; Wang, R.; Yang, P. R-TTWD: Robust device-free through-the-wall detection of moving human with WiFi. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 1090–1103. [CrossRef]
4. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. A survey on sensor networks. *IEEE Commun. Mag.* **2002**, *40*, 102–114. [CrossRef]
5. Patwari, N.; Ash, J.N.; Kyperountas, S.; Hero, A.O.; Moses, R.L.; Correal, N.S. Locating the nodes: Cooperative localization in wireless sensor networks. *IEEE Signal Processing Mag.* **2005**, *22*, 54–69. [CrossRef]
6. Xiao, F.; Sha, C.; Chen, L.; Sun, L.; Wang, R. Noise-tolerant localization from incomplete range measurements for wireless sensor networks. In Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM), Hong Kong, China, 26–30 April 2015; pp. 2794–2802.
7. Karbasi, A.; Oh, S. Robust localization from incomplete local information. *IEEE/ACM Trans. Netw.* **2013**, *21*, 1131–1144. [CrossRef]
8. Xiao, F.; Liu, W.; Li, Z.; Chen, L.; Wang, R. Noise-tolerant wireless sensor networks localization via multinorms regularized matrix completion. *IEEE Trans. Veh. Technol.* **2018**, *67*, 2409–2419. [CrossRef]
9. Maz'ya, V.; Schmidt, G. On approximate approximations using Gaussian kernels. *IMA J. Numer. Anal.* **1996**, *16*, 13–29. [CrossRef]
10. Meng, D.; De La Torre, F. Robust matrix factorization with unknown noise. In Proceedings of the 2013 IEEE International Conference on Computer Vision, Sydney, NSW, Australia, 1–8 December 2013; pp. 1337–1344.
11. Yao, J.; Cao, X.; Zhao, Q.; Meng, D.; Xu, Z. Robust subspace clustering via penalized mixture of Gaussians. *Neurocomputing* **2018**, *278*, 4–11. [CrossRef]
12. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–38. [CrossRef]
13. Shang, Y.; Ruml, W.; Zhang, Y.; Fromherz, M.P. Localization from mere connectivity. In Proceedings of the 4th ACM international symposium on Mobile ad hoc networking computing, Annapolis, MD, USA, 1–3 June 2003; pp. 201–212.
14. Sheng, X.; Hu, Y.H. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Signal Processing Mag.* **2005**, *53*, 44–53. [CrossRef]
15. Tomic, S.; Beko, M.; Dinis, R. RSS-based localization in wireless sensor networks using convex relaxation: Noncooperative and cooperative schemes. *IEEE Trans. Veh. Technol.* **2015**, *64*, 2037–2050. [CrossRef]
16. So, A.M.C.; Ye, Y. Theory of semidefinite programming for sensor network localization. *Math. Program.* **2007**, *109*, 367–384. [CrossRef]
17. Shamsi, D.; Taheri, N.; Zhu, Z.; Ye, Y. Conditions for correct sensor network localization using SDP relaxation. In *Discrete Geometry and Optimization*; Springer: Heidelberg, Germany, 2013; pp. 279–301.
18. Nguyen, T.L.N.; Shin, Y. Matrix completion optimization for localization in wireless sensor networks for intelligent IoT. *Sensors* **2016**, *16*, 722. [CrossRef] [PubMed]
19. Sun, X.; Chen, T.; Li, W.; Zheng, M. Perfomance research of improved mds-map algorithm in wireless sensor networks localization. In Proceedings of the 2012 International Conference on Computer Science and Electronics Engineering (ICCSEE), Hangzhou, China, 23–25 March 2012; pp. 587–590.
20. Feng, C.; Valaee, S.; Au, W.S.A.; Tan, Z. Localization of wireless sensors via nuclear norm for rank minimization. In Proceedings of the 2010 IEEE Global Telecommunications Conference GLOBECOM 2010, Miami, FL, USA, 6–10 December 2010; pp. 1–5.
21. Liu, C.; Shan, H.; Wang, B. Wireless Sensor Network Localization via Matrix Completion Based on Bregman Divergence. *Sensors* **2018**, *18*, 2974. [CrossRef] [PubMed]
22. Lee, D.D.; Seung, H.S. Algorithms for non-negative matrix factorization. *Adv. Neural Inf. Processing Syst.* **2001**, 556–562.
23. Sun, R.; Luo, Z.Q. Guaranteed matrix completion via non-convex factorization. *IEEE Trans. Inf. Theory* **2016**, *62*, 6535–6579. [CrossRef]
24. Koren, Y.; Bell, R.; Volinsky, C. Matrix factorization techniques for recommender systems. *Computer* **2009**, *8*, 30–37. [CrossRef]
25. Liu, H.; Wu, Z.; Li, X.; Cai, D.; Huang, T.S. Constrained nonnegative matrix factorization for image representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 1299–1311. [CrossRef] [PubMed]
26. Parikh, N.; Boyd, S. Proximal algorithms. *Found. Trends®Optim.* **2014**, *1*, 127–239. [CrossRef]

27. Chen, L.; Zhang, H.; Lu, J.; Thung, K.; Aibaidula, A.; Liu, L.; Chen, S.; Jin, L.; Wu, J.; Wang, Q.; Zhou, L.; Shen, D. Multi-label nonlinear matrix completion with transductive multi-task feature selection for joint MGMT and IDH1 status prediction of patient with high-grade gliomas. *IEEE Trans. Med. Imaging* **2018**, *37*, 1775–1787. [CrossRef] [PubMed]

28. Armijo, L. Minimization of functions having Lipschitz continuous first partial derivatives. *Pac. J. Math.* **1966**, *16*, 1–3. [CrossRef]

29. Chen, L.; Yang, G.; Chen, Z.; Xiao, F.; Shi, J. Correlation consistency constrained matrix completion for web service tag refinemen. *Neural Comput. Appl.* **2015**, *26*, 101–110. [CrossRef]

30. De La Torre, F.; Black, M.J. A framework for robust subspace learning. *Int. J. Comput. Vis.* **2003**, *54*, 117–142. [CrossRef]

31. Srebro, N.; Jaakkola, T. Weighted low-rank approximations. In Proceedings of the 20th International Conference on Machine Learning (ICML-03), Washington, DC, USA, 21–24 August 2003; pp. 720–727.

32. Buchanan, A.M.; Fitzgibbon, A.W. Damped newton algorithms for matrix factorization with missing data. In Proceedings of the 2005 Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–26 June 2005; pp. 316–322.

33. Keshavan, R.H.; Montanari, A.; Oh, S. Matrix completion from a few entries. *IEEE Trans. Inf. Theory* **2010**, *56*, 2980–2998. [CrossRef]